



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/522,753	03/10/2000	Ronald M. Evans	SALK1510-3	4924

30542 7590 05/23/2005

FOLEY & LARDNER
P.O. BOX 80278
SAN DIEGO, CA 92138-0278

EXAMINER

LEFFERS JR, GERALD G

ART UNIT	PAPER NUMBER
----------	--------------

1636

DATE MAILED: 05/23/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/522,753

Applicant(s)

EVANS ET AL.

Examiner

Gerald G. Leffers Jr., PhD

Art Unit

1636

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 15 February 2005.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 3-5, 9-10, 12-14, 16-25 & 38 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☒ Claim(s) 9, 10, 12, 13, 17 and 38 is/are allowed.
- 6) ☐ Claim(s) 4, 5, 14, 16, 18, 19, 21-25 is/are rejected.
- 7) ☒ Claim(s) 3 and 20 is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. _____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____

- 4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☒ Other: Exhibit A (Search Report SEQ ID No: 5)

DETAILED ACTION

Response to Amendment

Receipt is acknowledged of a supplementary response, filed 2/15/2005, in which the response explicitly answered each of the grounds of rejection made in the previous office action mailed on 4/20/2004. In the original response to the previous office action, filed on 8/12/2004, several claims were amended (claims 4, 5, 9, 12, 14, 23 and 25). Claims 3-5, 9-10, 12-14, 16-25 & 38 are pending and under consideration in the instant application.

Any rejection of record not addressed herein is withdrawn. This action is not final as there are new grounds of rejection presented herein that were not necessitated by applicants' amendment of the claims in the response filed on 8/12/2004.

Priority

Applicant has not complied with one or more conditions for receiving the benefit of an earlier filing date under 35 U.S.C. 120 as follows:

The later-filed application must be an application for a patent for an invention which is also disclosed in the prior application (the parent or original nonprovisional application or provisional application); the disclosure of the invention in the parent application and in the later-filed application must be sufficient to comply with the requirements of the first paragraph of 35 U.S.C. 112. See *Transco Products, Inc. v. Performance Contracting, Inc.*, 38 F.3d 551, 32 USPQ2d 1077 (Fed. Cir. 1994). The prior application to which the instant application seeks priority is U.S. Application Serial No. 08/522,726, filed 9/1/1995 (now U.S. Patent No. 6,489,441). The '726 application discloses only 3 sequences that correspond to SEQ ID NOS: 1-

Art Unit: 1636

3 of the instant application. Each of the pending claims is directed to an isolated polynucleotide that (i) has a recited percent identity to one of SEQ ID NOS: 4, 6, & 8; or (ii) encodes a polypeptide having a recited percent identity to one of SEQ ID NOS: 5, 7 & 9. The prior application does not disclose these particular sequences. Therefore, the prior application does not provide support for the broadly recited genus of polynucleotides encompassed by the pending claims. Accordingly, the priority date for the pending claims is the filing date of the instant application (3/10/2000).

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 4, 19, 21-22 are rejected under 35 U.S.C. 102(b) as being anticipated by Chen et al (Nature, October 1995, Vol. 377, No. 6548, pages 454-457; see the entire reference). **This is a new rejection.**

Chen et al teach the identification and characterization of a transcriptional co-repressor that is an SMRT (i.e. silencing mediator for retinoid and thyroid hormone receptors). The SMRT polypeptide taught by Chen et al is encoded by a polynucleotide sequence that encodes a polypeptide that is ~94% identical to the sequence of SEQ ID NO: 5 (see the attached Exhibit A,

Art Unit: 1636

result #6 for accession number HSU37146). Thus, the polynucleotide sequence taught by Chen et al anticipates the broad genus of polynucleotides encompassed by the instant claims.

Claims 4, 19, 21-22 are rejected under 35 U.S.C. 102(a) as being anticipated by Ordentlich et al (PNAS USA, 16 March 1999, Vol. 96, No. 6, pages 2639-2644; see the entire reference). **This is a new rejection.**

Ordentlich et al teach unique forms of the human and mouse nuclear receptor co-repressor SMRT. In particular, Ordentlich et al teach nucleic acids, described by accession numbers AF113003 & AF113001, that encode polypeptides with 100% & 88.2% identity with SEQ ID NO: 5, respectively (e.g. see results 1 & 3 of the search report provided as Exhibit A).

Claims 4, 19, 21-22 are rejected under 35 U.S.C. 102(a) as being anticipated by Park et al (PNAS USA, 30 March 1999, Vol. 96, No. 7, pages 3519-3524; see the entire reference). **This is a new rejection.**

Park et al teach the identification of an extended isoform of SMRT termed SMRTe by the authors. In particular, Park et al teach nucleic acids, described by accession numbers AF125672 & AF125671, that encode polypeptides with ~98% and ~82% identity with SEQ ID NO: 5 (e.g. see results 2 & 4 of the search report provided as Exhibit A).

Claim Rejections - 35 USC § 112

The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it

Art Unit: 1636

pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 5 & 18 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter that was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. **This is a new rejection.**

Claim 5 is directed to an isolated polynucleotide encoding an SMRT co-repressor, or a peptide portion thereof, where the SMRT co-repressor or peptide portion thereof is capable of mediating the transcriptional silencing of at least one member of the steroid/thyroid hormone superfamily of receptors. The claim then recites the limitation of "...wherein said SMRT co-repressor or peptide portion thereof is encoded by a polynucleotide having at least 80% sequence identity with SEQ ID NO: 4". This latter recitation appears to specify that the polynucleotide having at least 80% identity to SEQ ID NO: 4 is a different polynucleotide from the one that is actually claimed. As such, the nucleotide that is actually claimed can be any nucleotide that encodes the same SMRT protein, or portion thereof, that is encoded by the second nucleotide sequence (e.g. the first polynucleotide encodes the SMRT protein, or portion thereof, and which is different from the second polynucleotide because of the degeneracy of the genetic code and/or because the two different polynucleotides comprise additional, different sequences from one another). There does not appear to be support anywhere in the originally filed specification or claims for this formulation. Therefore, the phrase "...wherein said SMRT co-repressor or peptide portion thereof is encoded by a polynucleotide having at least 80% sequence identity with SEQ ID NO: 4" is impermissible NEW MATTER.

Art Unit: 1636

Claims 23-25 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter that was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

This is a new rejection necessitated by applicants' amendment of the claims in the response filed on 8/12/2004.

Claim 23 is directed to a genus of oligonucleotides that are identifiable under "suitable stringency conditions" with respect to other nucleic acid sequences. The term "suitable stringency conditions" is used in the context of an identified oligonucleotide comprising at least 15 nucleotides that hybridizes to a polynucleotide of claim 4, but not to a polynucleotide encoding SEQ ID NO: 11 or to a polynucleotide encoding an amino acid sequence consisting of amino acids 1031 to 2517 of SEQ ID NO: 5. The specification does not define these exact conditions and the skilled artisan has no basis to visualize what these "suitable" conditions might be. Similarly, claim 25 specifies that the oligonucleotide of claim 23 hybridizes under "suitable stringency conditions" to a polynucleotide encoding SEQ ID NO: 5 or SEQ ID NO: 7, but does not hybridize to a polynucleotide encoding SEQ ID NO: 9. Again, the exact hybridization conditions are not described. Thus, the rejected claims comprise a genus of oligonucleotides that must meet very particular hybridization requirements, yet there is no description of the hybridization conditions that will necessarily identify an oligonucleotide having the recited functional activity.

Art Unit: 1636

There does not appear to be any literal or implicit support in the originally filed claims or specification for claiming an isolated oligonucleotide comprising at least 15 nucleotides and having the particular hybridization characteristics recited in the rejected claims. Therefore, the rejected claims comprise impermissible NEW MATTER.

Further, there is no basis for the skilled artisan to predictably envision even a single oligonucleotide sequence that will meet all of the hybridization requirements recited in the rejected claims. Thus, the skilled artisan would not have been able to envision a sufficient number of embodiments to describe the claimed genus of oligonucleotides having very particular binding characteristics.

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter that the applicant regards as his invention.

Claims 5, 14, 16, 18 and 23-25 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. **These are new grounds of rejection.**

Claim 5 is vague and indefinite in that the metes and bounds of the phrase "...wherein said SMRT co-repressor or peptide portion thereof is encoded by a polynucleotide having at least 80% sequence identity with SEQ ID NO: 4" are unclear. Is the polynucleotide recited in this phrase necessarily the same as the isolated polynucleotide that is being claimed or can it be a second polynucleotide (e.g. a different polynucleotide that encodes the SMRT protein, or portion thereof, and which is different from the claimed polynucleotide because of the degeneracy of the

Art Unit: 1636

genetic code and/or because the two different polynucleotides comprise additional, different sequences from one another)?

Claim 14 recites a “first isolated polynucleotide encoding a SMRT co-repressor” and then recites that the first polynucleotide is selected from a Markush group of different polynucleotides. It is unclear how the nucleotides of part (d) can encode a SMRT co-repressor, or portion thereof, when they are *complementary* to sequences in parts (a)-(c) that actually do encode an SMRT co-repressor. Similarly, claim 16 also recites that the first polynucleotides of part (d) have 80% identity to the *complement* of sequences in parts (a)-(c).

Claim 23 is vague and indefinite in that the metes and bounds of the phrase “suitable stringency conditions” are unclear. The phrase is used in the context of an identified oligonucleotide comprising at least 15 nucleotides that hybridizes to a polynucleotide of claim 4, but not to a polynucleotide encoding SEQ ID NO: 11 or to a polynucleotide encoding an amino acid sequence consisting of amino acids 1031 to 2517 of SEQ ID NO: 5. The specification does not define these exact conditions and the skilled artisan has no basis to visualize what these “suitable” conditions might be.

Similarly, claim 25 specifies that the oligonucleotide of claim 23 hybridizes under “suitable stringency conditions” to a polynucleotide encoding SEQ ID NO: 5 or SEQ ID NO: 7, but does not hybridize to a polynucleotide encoding SEQ ID NO: 9.


Art Unit: 1636

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Gerald G. Leffers Jr., PhD whose telephone number is (571) 272-0772. The examiner can normally be reached on 6:30-4:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Remy Yucel can be reached on (571) 272-0781. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).


Gerald G Leffers Jr., PhD
Primary Examiner
Art Unit 1636
GERRY LEFFERS
PRIMARY EXAMINER

ggl

Mon Apr 19 08:15:02 2004

us-09-522-753-5.ige

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2004, 22:23:25 ; Search time 20448 Seconds
(without alignments)
5335.212 Million cell updates/sec

Title: US-09-522-753-5
Perfect score: 13215
Sequence: 1 MSGSTQLVAQWTRATEPRYP.....WDBEPKPLLCQYETLSDE 2517

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09522753/runat_15042004_143737_17428/app_query.fasta_1.2695
-DB=GenEmbl -OPT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=100 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09522753@cgn_1_14601@runat_15042004_143737_17428 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOC=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:
2: gb_hgt:
3: gb_in:
4: gb_in:
5: gb_ov:
6: gb_ov:
7: gb_ph:
8: gb_ph:
9: gb_pr:
10: gb_ro:
11: gb_ro:
12: gb_sy:
13: gb_sy:
14: gb_sy:
15: gb_sy:
16: gb_sy:
17: gb_sy:
18: gb_sy:
19: gb_sy:
20: gb_sy:
21: gb_sy:
22: gb_sy:
23: gb_sy:
24: gb_sy:
25: gb_sy:
26: gb_sy:
27: gb_sy:
28: gb_sy:

Exhibit A
09/522,753 Page 1
Applicant Copy

29: em_v2:
30: em_hgt_hum:
31: em_hgt_inv:
32: em_hgt_mus:
33: em_hgt_mus:
34: em_hgt_mus:
35: em_hgt_mus:
36: em_hgt_mus:
37: em_hgt_mus:
38: em_hgt_mus:
39: em_hgt_mus:
40: em_hgt_mus:
41: em_hgt_mus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	13215	100.0	8561	9	AF113003	AF113003 Homo sapi
2	12978	98.2	8561	9	AF113003	AF113003 Homo sapi
3	11067	83.7	8338	10	AF113001	AF113001 Mus muscu
4	10832	82.0	8561	10	AF113001	AF113001 Mus muscu
5	9894	74.9	7445	10	AF113002	AF113002 Mus muscu
6	7885	59.7	5339	9	HSU37146	U37146 Human silen
7	4612	34.9	2330	9	S83390	S83390 T3 receptor
8	4194	31.7	7839	9	AB028970	AB028970 Homo sapi
9	4187	31.7	7839	10	MMU35312	U35312 Mus musculu
10	4147.5	31.4	7940	6	AX578069	AX578069 Sequence
11	4147.5	31.4	7940	9	AF044209	AF044209 Homo sapi
12	4043	30.6	8088	9	AF087856	AF087856 Homo sapi
13	4009.5	30.3	8539	5	AF495886	AF495886 Xenopus l
14	3736.5	28.3	194840	2	AC139377	AC139377 Mus muscu
15	3479.5	26.3	254430	2	AC139377	AC139377 Mus muscu
16	3445	26.1	218119	2	AC136560	AC136560 Rattus no
17	3344.5	25.3	213339	2	AC121005	AC121005 Rattus no
18	3178	24.0	2842	9	BC004326	BC004326 Homo sapi
19	3012	22.8	2934	10	BC047524	BC047524 Homo sapi
20	2891.5	21.9	205233	9	AC073916	AC073916 Homo sapi
21	2187	16.5	31320	9	AF303586	AF303586 Homo sapi
22	2155	16.3	1937	5	BC054296	BC054296 Xenopus l
23	2091.5	15.8	6338	10	AB093281	AB093281 Mus muscu
24	1900	14.4	6531	9	AB019524	AB019524 Homo sapi
25	1839	13.9	4235	10	MUSRIPI13	L78294 Mus musculu
26	1687	12.8	1831	9	BC050594	BC050594 Homo sapi
27	1633	12.6	1831	9	BC050594	BC050594 Homo sapi
28	1633	12.4	1731	9	BC058511	BC058511 Homo sapi
29	1628.5	12.3	3035	5	BC056862	BC056862 Homo sapi
30	1396	10.6	161970	2	AC027706	AC027706 Homo sapi
31	1391	10.5	2934	9	AK127788	AK127788 Homo sapi
32	1272	9.6	3575	9	HSN803410	HSN803410 Homo sapi
33	1172.5	8.9	3937	9	AF303585	AF303585 Homo sapi
34	1162	8.8	133347	5	AL590153	AL590153 Zebraphis
35	1113	8.4	630	6	AX677743	AX677743 Sequence
36	1099	8.3	11236	3	AF175223	AF175223 Drosophi
37	1076	8.1	802	6	BD221548	BD221548 Human gen
38	960	7.3	525	6	AX677866	AX677866 Sequence
39	932	7.1	530	9	HSU80750	U80750 Homo sapien
40	932	7.1	530	9	HSU80750	U80750 Homo sapien
41	880	6.7	738	6	AX753058	AX753058 Sequence
42	845	6.4	738	6	BD021376	BD021376 Novel gen
43	845	6.4	738	6	BD101314	BD101314 Novel gen
44	774.5	5.9	673	6	AX677865	AX677865 Sequence
45	763.5	5.8	135351	9	AC068279	AC068279 Homo sapi
46	760.5	5.8	150739	2	AC130376	AC130376 Homo sapi
47	750	5.7	1793	6	AX057740	AX057740 Homo sapi
48	746.5	5.6	520	6	AX396270	AX396270 Sequence
49	743	5.6	40871	2	AC020019	AC020019 Drosophi
50	743	5.6	160440	3	AC023741	AC023741 Drosophi

C 51	743	5.6	187921	3	AC104627	AC104627 Drosophil
C 52	743	5.6	302665	3	AB003490	Downes,M.R., Odentlich,P. and Evans,R.M.
C 53	736	5.6	165791	9	AC127533	Direct Submission
C 54	730	5.5	299925	1	AP005045	Submitted (11 DEC-1998) Gene Expression Laboratory, The Salk
C 55	726.5	5.5	33529	6	AB166425	Institute for Biological Studies, 10010 Torrey Pines Road, La
C 56	723	5.5	193894	9	AC14730	Jolla, CA 92037, USA
C 57	719	5.4	2036	10	MM022016	Location/Qualifiers
C 58	717.5	5.4	164485	2	AC026106	1. .561
C 59	716.5	5.4	339650	1	AC039108	/organism="Homo sapiens"
C 60	715.5	5.4	180591	9	AC106029	/mol_type="mRNA"
C 61	714.5	5.4	1599	10	AF124821	/db_xref="taxon:9606"
C 62	709.5	5.4	192181	9	AC133644	/chromosome="12"
C 63	708.5	5.4	320050	1	EX248336	/map="12q24"
C 64	708	5.4	200389	9	AC134678	/tissue type="pituitary"
C 65	704	5.3	321250	1	AC039111	2. .755
C 66	701	5.3	58343	1	AB089954	/function="transcriptional co-repressor"
C 67	700.5	5.3	164485	2	AC026106	/note="5SMT alpha; longer isoform than previously
C 68	700	5.3	13416	6	AX697987	repeated
C 69	700	5.3	60196	6	AX697987	/coding_start=1
C 70	696.5	5.3	204006	2	AC134881	/product="silencing mediator of retinoic acid and thyroid
C 71	691.5	5.2	7564	14	GFL309022	hormone receptor alpha"
C 72	691.5	5.2	123288	9	AP000547	/protein_id="AAD20946.1"
C 73	689	5.2	427	6	AX072504	/db_xref="GI:4454552"
C 74	689	5.2	69301	1	STU82965	/translation="MSGSTOLVAQTRATEPRYPHSLSYVQIARTDVLGLEYQH
C 75	689	5.2	348564	1	EX842574	HGRDYSLSHSGSIIOQRPSLSSEFOGNERSOELHLRPSHSYVLPGLKSEMEF
C 76	685	5.2	30690	6	E38020	ISKRELLLELDPLDPLATGPPAGSDELTKRSLLTKGLKLEPPSPPHPTPEL
C 77	685	5.2	30690	6	BD097649	ELVPELLEKESLIDMDRDVREITVYEQISKLKKQOQLEERAKPPSPKPPVPP
C 78	685	5.2	64957	1	AB032367	ISKLSLVQIIVDENKKAABAHRILEGLPVELLYNPQDPTQYHNIKNQM
C 79	685	5.2	302675	1	AP005024	RKLLIDYFKRNHARKKQKFCQSDQMLKSKVERIENRPRRAKESKRYEYE
C 80	684	5.2	16188	1	AB069698	KKFFERKQRLQRMOSVQSGLSMSARSSEIIDLSEQENLEKQKQL
C 81	683.5	5.2	155343	9	AC087386	AVIPELDADQKIFINNGMLADPMKYQKQVQRMVMSSEKEKTFKKMQHKN
C 82	681.5	5.2	196492	5	AC139725	FGLLIEFLERTACVLYLTKNNKLSLVRSVRRKSKSQQQQQQQQQQQQ
C 83	677.5	5.1	299050	1	SC039119	QPMFRSQQEKKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKEKE
C 84	673.5	5.1	15397	1	AB093554	KGRKTSQGEKRRKIRTSMAANESEAIPTQSAELASLMDNSETSEWTEETAK
C 85	669	5.1	75236	6	AX600586	KGLLEGRNWAIAVMVSTVSCQNFVYKKNQDLDELAQHKLVKKEKERNARK
C 86	666.5	5.0	305550	1	EX248342	KKKAPAAEAPVVEDEMEASGVNVEEAEALHAGNVEVRCSCGPA
C 87	666	5.0	74457	1	AF210249	TVNSSEDTESIPSPHTEAAKQNGKVPATLADGPPGPPPTPRRTSRAPISPT
C 88	665	5.0	14860	1	AB007093	ASENLTAPPPAPSPAPVVPKEKEKETAAPVVEGEQKPPAAEALVDVTG
C 89	665	5.0	299925	1	AP005045	KAEELKSECTEAEAGPKDAENATAGALKAEKKEGSGRATLTKSGAPQD
C 90	665	5.0	346051	15	EX842580	SUSNLSADSVBAGGDKNLLSPRSLTLTGDPRANASQKLDLKOIKQRAA
C 91	664.5	5.0	96742	1	AY211023	IPPIQTKVTHSPEDDAETPAPPAFPFPPVPPVPPVPPVPPVPPVPPVPPVPP
C 92	662.5	5.0	138203	1	AY130323	DEKAPAAQKLPDGPDPCTGGLFPFPPVPPVPPVPPVPPVPPVPPVPPVPP
C 93	662	5.0	135301	14	BHVLCGEN	HTDAEVLPRPTTSPPLSISSAKSVLESQIGASQGSVOLHVPESEKAPVG
C 94	661.5	5.0	43280	1	SFU78289	PVTMCEPLPMPKCLAPSGVQKQESPRGQGPPELPGVTAQASVLRGALGSPV
C 95	658.5	5.0	15311	1	AF007164	GGSIKGPISRVFSDSAITRGSITGHPADVLYKGTITRIIGEDSPRLDRGDS
C 96	654.5	5.0	113193	6	AF357202	LPKGMIVYEGKKHVLSEGGMSVTCQSKEDGRSSGPPHETAAKRTYDMMEGRVR
C 97	654.5	5.0	113193	6	AF357202	AISRAIEGLMGRAIPPEHSPHLEKHIRGSIQICIPRVSVEAQEYLRREKLL
C 98	654	4.9	22449	1	SFSNECDE	KRBGPPPPPSRDLTAYKQALGPLKHPAHEGLVATVKEAGRSIHEIPRELRHT
C 99	654	4.9	22449	1	SFSNECDE	PELPAHPLKSGSITQGTPLKYDTGASTGSKHVRSLIGSGRTPFPVHPDVA
C 100	654	4.9	313800	1	SC039114	DARALRACYSLESLKSPGASSSGSIARGAPVIVPELKGPRQSPITVEDHAPAG
						HLPRCEPVMEPTFRLQEGSLSSKASQDKLTSTPREIAKSPHSTVPHHPHPSIP
						YELLHGVMSVDLIRSHIPLAFDTISIRGIFLDAAYAYILPHILAFNETHLYPPY
						LINGEDYDALENQTIINDYITTSQMHNTATANAQMLAGLSFRESSLALNYAA
						GPGRKDLISQVPLVPLVPTGTPATMDRLAYLTPAPQFSSHSSPLPGPPTH
						LTKPTSTSSRERDRDRDREREKSLTSTTVEHAPIMRPTFESSSSSSSSSG
						GGGSSSRPASHSHAHQHSIPRTQDALQQRPSVLNTHMGKIIITAVFSPKTVLRS
						TSTSPVPAATFPFATHCPGLGTLGVYPTLMEEVLLPKEAPVAPRPRPADTGA
						FLAKBARSGLPEASPSKSGEPRLVPPVSGHATIAIPAKNALPHASPPKPPAPPA
						SELDNHEHLEGRKPKQPVKLGGEAAHPLRPLPESQSPSSPSLQTPAGVKGHVR
						RAEDHREKTSQKPSIOELERSLGHGSSYSGPESVPSVSSSLQTPAGVKGHVR
						VTLACHISEVITQDTRHPQQLSAPLPAPYSPGASCVLDLRRPDDVLPDPPH
						GAPAEPSHSGQSGSPENKTSVLGGEDGIERVSPPEGMTGHRSAVPLVLYRD
						GEOTEPSRMGSKSFONTSPAPFSLKTESNMMVKSKOINKLNTNRNPEVNI
						SQPEIFIMFAITGLMTIRSQVQEHASTNMGLEAIKIKALMKYQWSESPLS
						ANAFPLNALSALPAMPTAADGRSDHITSPGGGKAKYSGRPSRKAKGAFGLA
						SGDRPESSVSHSGDCDPLTNVWEDRPSAGSTPPFPIPLMLRQAGVMAFPP
						PPCLAGSGGLAGPHAWDEEPKLLCSQYETLSDS8"

ALIGNMENTS

RESULT 1	AF113003	8561 bp	mRNA	linear	PRI 20-MAR-1999
LOCUS	AF113003				
DEFINITION	AF113003				
ACCESSION	AF113003				
VERSION	AF113003.1	GI:4454551			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					

ORIGIN

Alignment Scores:

Pred. No.: 26-259

Score: 13015.00

Percent Similarity: 100.00%

Matches: 2517

Conservative: 0

Length: 8561

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-522-753-5 (1-2517) x AF113003 (1-8561)

QY	1	MetSerGlnSerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro	20
DB	2	ARGTCGGCTCCACACAGCTTGTGGCRACGACGTGGAGGCCACTGAGCCCGCTACCCG	61
QY	21	ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu	40
DB	62	CCCAACAGCTTTCATCCACGATCGCCGAGATCGCCGAGCAGCACAGGAGCTGGGCTCCTG	121
QY	41	GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln	60
DB	122	GAGTACACGACCACTCCCGGACTATGCTCTCCCACTGTTCGGCGGGCTCCATCATCCAG	181
QY	61	ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln	80
DB	182	CCCAAGCGCGAGGCGCTCCCTGCTGTGAGTTCAGCCCGGGAATGAAACGGTCCCCAG	241
QY	81	GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet	100
DB	242	GAGTCTCACCTCGGGCCAGAGTCCCACTCATCTGCGGAGCTGGGGAAGTCAGAGATG	301
QY	101	GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro	120
DB	302	GAGTTCATTGAAGCAAGCGCTTCGGCTAGAGCTGCTGCTGACCCCTGCTGGACCG	361
QY	121	SerProLeuLeuAlaThrGlnProAlaGlySerGluAspLeuThrLysAspArgSer	140
DB	362	TACCCCTGTGCGCCACGGCCAGCTTGGGATCTGAAGACTTCAACAGGACCGTAGC	421
QY	141	LeuThrGlyLysLeuGluProValSerProSerProHisThrAspProGluLeu	160
DB	422	CTGACGGGCAAGCTGGACCGGTGTCTCCCGCCAGCCCGCCACACTGACCCCTGAGCTG	481
QY	161	GluLeuValProProArgLeuSerLysGluLeuIleGlnAsnMetAspArgValAsp	180
DB	482	GAGCTGTGTGGCCACCGCTGTCCAGGAGGAGCTGATCCAGAACATGGACCGGTGGAC	541
QY	181	ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysGlnGlnLeu	200
DB	542	CGAGATACCACTGATGAGCAGCATCTTATGCTGAAGAGAGGAGGAGGAGGAGGAGG	601
QY	201	GluGluAlaAlaLysProGluProGluLysProValSerProProIleGlu	220
DB	602	GAGGAGGAGGCTGCCAAGCGCCGAGCTGAGAGCCCGTGTCAACCGCCGCAATCGAG	661
QY	221	SerLysHisArgSerLeuValGlnIleTyrAspGluAsnArgLysAlaGluAla	240
DB	662	TCGAAGCACCGCAGCTGTGTGAGATCATCTAGACAGAGACCCGGAAGAGGTGAGCT	721
QY	241	AlaHisArgIleLeuGluGlyProGlnValGluLeuProLeuTyrAsnGlnPro	260
DB	722	GCAATCGGATTCGGAAGGCTGGGGCCCGGTGGAGCTGCGCTGTACAACACGCCC	781
QY	261	SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu	280
DB	782	TCCGACACCCCGGAGTATCATGAGAAATCAAAATAAACAGGCGATGCGGAAGAGCTA	841
QY	281	IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln	300
DB	842	ATCTTGACTTCAAGGAGGAGGATCAGCTCGGAACATATGAGAGCAAGTCTTGCCAG	901
QY	301	ArgTyrAspGlnLeuMetGluAlaLeuGlyLysValGluArgIleGluAsnAsnPro	320
DB	902	CGCTATGACAGCTCATGAGGCGCTTGGAAAAAAGGTGGAGCGCATCGAAAAACACCG	961
QY	321	ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle	340
DB	962	CGCCGCGCGGCCAAGAGAGCAGAGGTGCGGAGTACTACGAAGAGCAAGTCTCCCTGAGATC	1021

QY	341	ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly	360
DB	1022	CGCAAGCAGCGGAGCTGCAGGAGCGCATGCAGAGCGGGTGGCCAGCGGGCGAGTGG	1081
QY	361	LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleIleIleIleSer	380
DB	1082	CTGTCTCATGTGCGCGCCGCGCAGCAGCAGAGGTGTGAGATCATCATGAGCTCTCA	1141
QY	381	GluGlnGluAspLeuGluLysGlnMetArgGlnLeuAlaValIleProMetLeuTyr	400
DB	1142	GAGCAGGAGAACTGGAGAGCAGATGCGCCAGCTGGCTGTATGCGCCGCCATGCTGAC	1201
QY	401	AspAlaAspGlnArgIleLysPheIleAsnMetGlnGlyLeuMetAlaAspProMet	420
DB	1202	GAGCTGACCACTGAGCGCATCAAGTTCATCAACATGACCGGCTTATGCGCCGCCATG	1261
QY	421	LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGlyLysIleThrPhe	440
DB	1262	AAGGTGTACAAAGACCCCGCAGTGCATGAACATGTGAGTGCAGCGAGAGAGAGACCTTC	1321
QY	441	ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg	460
DB	1322	CGGAGAGAGTTCATGCAGCATCCCAAGAACTTTGGCTGATCGCATCATTCCTCGAGAGG	1381
QY	461	LysThrValAlaGluCysValLeuTyrTyrLeuThrLysLysAsnGluAsnTyrLys	480
DB	1382	AGACAGTGGCTGAGTGGCTCTCTATTACTACTGACTAAGAGAGAGTATGAGTATGAG	1441
QY	481	SerLeuValArgSerTyrArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln	500
DB	1442	AGCTTGGTGAAGAGAGCTATCGCGCCGCGGCAAGAGCAGCAGCAACAGCAGCAG	1501
QY	501	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	520
DB	1502	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1561
QY	521	AspGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGlu	540
DB	1562	GATGAGAGGAGTGAAGAAAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG	1621
QY	541	AspLysGluAsnLeuLeuLysThrAspThrSerGlyGluAsnAspGlu	560
DB	1622	GACAGGAGAGCTTCCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1681
QY	561	LysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgLysGly	580
DB	1682	AAGGAGGCTGTGCTCTCAAGCGCAAACTGCCAACAGCAGCAGGAGAGAGAGAG	1741
QY	581	ArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluAlaIleThrProGlnGln	600
DB	1742	CGCATCACCCGCTCATGTGCTANTGAGGCAACAGAGAGAGAGAGAGAGAGAGAG	1801
QY	601	SerAlaGluLeuAlaSerMetGluLeuAsnGlnSerSerArgTyrThrGluGluMet	620
DB	1802	AGCGCCGAGCTGCTCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1861
QY	621	GluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArgMet	640
DB	1862	GAACAGCAACCAAGGTCTCTGGAAACAGCGCGCACTGGTGGCCATCGCCCGATG	1921
QY	641	ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysValArgGln	660
DB	1922	GTGGGCTTCCAGCTGTGTGCGCATGTGAGAACTTCTACTTCACTACAGAGAGAG	1981
QY	661	AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg	680
DB	1982	AACTTCGATGATCTTTCAGCAGCAGCAAGCTGAAGATGGAGAGAGAGAGAGAG	2041
QY	681	ArgLysLysLysLysAlaProAlaAlaAsnGluGluAlaAlaPheProProValVal	700
DB	2042	AGGAAGAAGAAACAAAGCGCGCGCGCCAGCAGGAGAGGCTGCATTCCCGCCGCTG	2101

Qy 701 GluAspGluMetGluAlaSerGlyValSerGlyAsnGluGluMetValGluGlu 720
Db 2102 GAGGATGAGGAGTGGAGGCTCGGGCTGAGCGAAATGAGGAGGAGATGCTGGAGGAG 2161
Qy 721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
Db 2162 GCTGAAGCTTACATGCTCTGGGATGAGGTGCCAGAGGGGAAATGCTGGCCAGCC 2221
Qy 741 ThrValAsnAsnSerAspThrGluSerIleProSerProHisThrGluAlaAlaLys 760
Db 2222 ACTGTCAACACAGCTCAGACCGAGAGCATCCCTCTCTCCACACTGAGGCCGCCAAG 2281
Qy 761 AspThrGlyGlnAsnGlyProLysProAlaThrLeuGlyAlaAspGlyProProPro 780
Db 2282 GACACAGGCGAGATGGGCCCAAGCCCCCAGCCACCTTGGGGCGCGAGCGGCCACCCCA 2341
Qy 781 GlyProThrProProArgArgThrSerArgAlaProIleGluProThrProAlaSer 800
Db 2342 GGCCTCACCCACCCACACGAGAGACATCCCGGGCCCCCATTTAGAGCCACCCCGGCTCT 2401
Qy 801 GluAlaThrGlyAlaProThrProProAlaProProSerProSerAlaProProPro 820
Db 2402 GAAGCCACCGAGGCCCTTACGCCCCACACGAGCACCCCTCGGCCCTCTGCACCTCTCT 2461
Qy 821 ValValProLysGluGluGlyGluGluGluThrAlaAlaAlaProValGluGluGly 840
Db 2462 GTGGTCCCNAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2521
Qy 841 GluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGlu 860
Db 2522 GAGGAGCAAGACCCCGCGCTGAGGAGCTGGCAGTGGACACAGGAGAGCGCGAGGAG 2581
Qy 861 ProValLysSerGluCysThrGluGluAlaGluGluGluProAlaLysGlyLysAla 880
Db 2582 CCGCTCAAGAGCGAGTGCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2641
Qy 881 GluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlySer 900
Db 2642 GAGCGCGCTGAGGCGACCGCGAGGGGGGCGCTCAAGCGCAGAGGAGGAGGAGGAGGAG 2701
Qy 901 GlyArgAlaThrAlaLysSerSerGlyAlaProGlnAspSerAspSerAlaThr 920
Db 2702 GGCAGGGCCACCATGCGCAAGAGCTCGGGCGCCCCCAGGACAGCGACCTCCAGTGCTACC 2761
Qy 921 CysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerPro 940
Db 2762 TGCAGTGCAGACGAGGTGGATGAGGCGGAGGCGGCGACAGAACCGGTGTGTCCCA 2821
Qy 941 ArgProSerLeuLeuThrProThrGlyAspProArgAlaAlaAlaSerProGlnLysPro 960
Db 2822 AGCCGCGCTCTTACCCCGATGGCGACCCCGGGCGCATGCCCTTCAACCCAGAGGAGCA 2881
Qy 961 LeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProIleGlnValThrLys 980
Db 2882 CTGGACCTGAAGCAGCTGAAGCAGCGAGCGGTGCGCATCCCGCCCCATCCAGGTCAACAAA 2941
Qy 981 ValHisGluProProArgGluAspAlaAlaProThrLysProAlaProProAlaProPro 1000
Db 2942 GTCCATGAGCCCCCGGGAGGAGCGAGCTCCACCAAGCCAGCTCCCGCCAGCCACCG 3001
Qy 1001 ProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSerProArg 1020
Db 3002 CCACCGCAAAACCTGACGCGGAGAGCGAGCGCTTCAAGCAGCTGGCAGGAGCGCCCGG 3061
Qy 1021 GlyLysSerArgSerProAlaProProAlaAspLysGluAlaPheAlaAlaGluAlaGln 1040
Db 3062 GGCAGAGCAGGAGCGCGCAGCCCGCCGCGACAGAGGAGGCTTCGCGAGCGCGGCGCAG 3121
Qy 1041 LysLeuProGlyAspProProCysTrrThrSerGlyLeuProPheProValProProArg 1060
Db 3122 AAGTGTCTGGGAGCCCCCTTGTGAGCTTCCGGGCTGCGCTTCCCGGCTGCGCCCGCT 3181
Qy 1061 GluValLysLysAlaSerProHisAlaProAspProSerAlaPheSerTyrrAlaProPro 1080

Db 3182 GAGGTGATCAAGSCCTCCCGCATGCCCGAGACCCCTCAGCGCTTCTCTACGCTCCACCT 3241
Qy 1081 GlyHisProLysProLeuGlyLeuHisAspThrAlaAspProValLeuProAspProPro 1100
Db 3242 GGTCAACCCACTCCCTCGGCTCCATGACACTGCCCGCGCTGCTGCGCGCGCCACCC 3301
Qy 1101 ThrLysSerAspProProProLeuLysSerSerAlaLysHisProSerValLeuGluArg 1120
Db 3302 ACCATCTCCAAACCGCCCTCCCTCATCTCTGCAAGACACCCCGAGGCTCTCGAGAG 3361
Qy 1121 GlnIleGlyValIleSerGlnGlyMetSerValGlnLeuHisValProTyrrSerGluHis 1140
Db 3362 CAATAGGTGCATCTCCCAAGGAATGTGGTCCAGTCCAGCTCCGCTCTCAGAGCAT 3421
Qy 1141 AlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLys 1160
Db 3422 GCCAAGCCCGCTGGGCTCTGTCCATGGGCTGCCCCCTGCTCATGACCCCAAAAG 3481
Qy 1161 LeuAlaProPheSerGlyValLysGlnGluLeuSerProArgGlyGlnAlaGlyPro 1180
Db 3482 CTGGCACCTTTGAGCGGAGTGAAGCAGGAGAGCTGTCCCAACCGGCGCAGGCTGGGCCA 3541
Qy 1181 ProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyThrAlaLeu 1200
Db 3542 CCGGAGAGGCTTGGGTGCCACAGCCAGAGGCGCTCCGTCTGAGAGGAGGAGGAGCTCTG 3601
Qy 1201 GlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgValProSerAsp 1220
Db 3602 GGTCTAGTTCCTGGCGGAGAGCATCCAAAGGATTTCCAGACACAGGCTGCTCGGAC 3661
Qy 1221 SerAlaIleThrTyrrArgGlySerIleThrHisGlyThrProAlaAspValLeuTyrrLys 1240
Db 3662 AGCCCATCACCATTACCGGGCTCCATCACCACGCGAGCTGAGCTGCTGTCAAG 3721
Qy 1241 GlyThrIleThrArgIleGlyGluAspSerProSerArgLeuAspArgGlyArgGlu 1260
Db 3722 GGCACCATCAGCAGGATCATCGGCGAGGACAGCGCGAGTCTGGACCGCGCGGAG 3781
Qy 1261 AspSerLeuProLysGlyHisValIleTyrrGluGlyLysGlyHisValLeuSerTyrr 1280
Db 3782 GACAGCTGCGCGAGGGCCAGCTCATCTAGAGGCAAGAGGCGACGCTTGTCTTAT 3841
Qy 1281 GluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerGlyPro 1300
Db 3842 GAGGCTGGCATCTCTGTGACCGAGTCTCCAGAGGAGCGCGAGAGGAGCTCAGGACCC 3901
Qy 1301 ProHisGluThrAlaAlaProLysArgThrTyrrAspMetMetGluGlyArgValGlyArg 1320
Db 3902 CCCCNTGAGACCGCGCCCGCCCGCCAGCGCATATGACATGATGAGGCGCGCTGGGAGCA 3961
Qy 1321 AlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProProGluArgHis 1340
Db 3962 GCCATCTCTCAGCAGCATCGAAGGTCTCATGGGCGGTGCTCCCGCGGAGCGACAC 4021
Qy 1341 SerProHisHisLeuLysGluGlnHisIleArgGlySerIleThrGlnGlyIlePro 1360
Db 4022 AGCCCCCACCACCTCAAGAGCAGCACCATCCGCGGTCCATCACCAAGGGATCCCT 4081
Qy 1361 ArgSerTyrrValGluAlaGlnGluAspTyrrLeuArgArgGluAlaLysLeuLysArg 1380
Db 4082 CGGTCTTACGTCAGGCAAGAGGACTCTCTGCTGGGAGGCGACAGCTCTTAAGCGG 4141
Qy 1381 GluGlyThrProProProProProSerArgAspLeuThrGluAlaTyrrLysThrGln 1400
Db 4142 GAGGCGACCGCTCCCGCCCAACCGCCCTCAGCGGACCTGACCGAGGCTCTCAAGAGCGAG 4201
Qy 1401 AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLeuValAlaThrValLysGlu 1420
Db 4202 GCGTGGCGCCCTGAAGCTCAAGCGCGCCATAGGGCTGTGTGGCCACCGGTGAAGGAG 4261
Qy 1421 AlaGlyArgSerIleHisGluIleProArgGluLeuArgHisThrProGluLeuPro 1440

Db 4262 GCGGGCGGCTCCATCATATGATCCCGCGGAGGAGCTGCGGCACACGCCGAGCTGCC 4321
Qy 1441 LeuAlaProArgProLeuLysGluGlySerIleThrGlnGlyThrProLeuLysTyrAsp 1460
Db 4322 CTGGCGCGCGCGCTCAAGGAGGGCTCCATCAGCGAGGACACCGCGCTCAAGTACGAC 4381
Qy 1461 ThrGlyAlaSerThrThrGlySerLysValHisAspValArgSerLeuIleGlySerPro 1480
Db 4382 ACCGGGGGTTCACACTGGCTCCAAAAGACACGAGTACGCTCCCTCATCGGAGCCCC 4441
Qy 1481 GlyArgThrPheProProValHisProLeuAspValMetAlaAspAlaArgAlaLeuGlu 1500
Db 4442 GCGCGGAGCTTCCACCGCTGCACCGCTGGATGTGATGCGCGACCGCGGCGCATGGAA 4501
Qy 1501 ArgAlaCysTyrGluGluSerLeuLysSerArgProGlyThrAlaSerSerSerGlyGly 1520
Db 4502 CGTGCTGTCTACGAGGAGAGCTTGAAGAGCGCGCCAGGACCGCCAGCAGCTCGGGGGC 4561
Qy 1521 SerIleAlaArgGlyAlaProValIleValProGluLeuGlyLysProArgGlnSerPro 1540
Db 4562 TCATTTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4621
Qy 1541 LeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLeuProArgLysSerProVal 1560
Db 4622 CTGACCTATGAGGACACACCGGGCACCTTTGCGGGCACCTCCACGAGGTTCGCGCGTG 4681
Qy 1561 ThrMetArgGluProThrProArgLeuGlnGlySerLeuSerSerSerLysAlaSer 1580
Db 4682 ACCATCGGGAGCCACCGCGCTGTCAGAGGGGCGAGCTTTCTGTCAGCAGGAGCATCC 4741
Qy 1581 GlnAspArgLysLeuThrSerThrProArgGluIleAlaLysSerProHisSerThrVal 1600
Db 4742 CAGGACCGAAGCTGACGTGACGCTGTGAGATGCCAAGTCCCGCGCACAGCAGCCGTG 4801
Qy 1601 ProGluHisHisProHisProLysSerProTyrGluHisLeuLeuArgGlyValSerGly 1620
Db 4802 CCGGAGCACACCCACACCTCTCGCTTATGAGACACCTGCTTCGGGGGTGAGTGGC 4861
Qy 1621 ValAspLeuTyrArgSerHisLeuProLeuAlaPheAspProThrSerIleProArgGly 1640
Db 4862 GTGAGCTGTATCGACGACATCCCTCGGCTTCGACCCCTCCATATCCCGCGCGC 4921
Qy 1641 IleProLeuAspAlaAlaAlaTyrTyrLeuProArgHisLeuAlaProAsnProThr 1660
Db 4922 ATCCCTCTGGAGCGAGCGCTGCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4981
Qy 1661 TyrProHisLeuTyrProTyrLeuIleArgGlyTyrProAspThrAlaAlaLeuGlu 1680
Db 4982 TACCGGCACTGTACCCACCTACCTATCCGCGGCTACCCCGACACGCGCGCTGGAG 5041
Qy 1681 AsnArgGlnThrIleIleAsnAspTyrIleThrSerGlnGlnMetHisAsnThrAla 1700
Db 5042 AACCGGAGACCATCATCATCATCATCATCTCGGAGGATGACCAACACACCGGCC 5101
Qy 1701 ThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSerProArgLysSerLeu 1720
Db 5102 ACCGCGATGGGCGAGCGAGTGTATGCTGAGGGGCGCTCTCGCGCGCGAGTCTCGCTG 5161
Qy 1721 AlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAspLeuSerGlnValProHisLeu 1740
Db 5162 GCACCTCACTAGCTGCGGGTCCCGAGGACATCATCGCTGCTCCAGTGCACACCTG 5221
Qy 1741 ProValLeuValProProThrProGlyThrProAlaThrAlaMetAspArgLeuAlaTyr 1760
Db 5222 CTGTGCTGTGGGCGCGCGACACCGAGGACCCCGACCGACCGACCGCTGCTGCTGCTGCT 5281
Qy 1761 LeuProThrAlaProGlnProPheSerSerArgHisSerSerSerProLeuSerProGly 1780
Db 5282 CTCCCGACCGCGCGCGCGCTTTCAGCAGCGCGCACAGCAGCTTCGCCACCTCTCCCGC 5341
Qy 1781 GlyProThrHisLeuThrLysProThrThrThrSerSerSerGluArgGluArgAspArg 1800
Db 5342 GGTCCACACACTTTCACAAAACCCACCGCTCTCTGTCGAGCGGGGAGACCGG 5401

Qy 1801 AspArgGluArgAspArgAspArgGluArgGluLysSerIleLeuThrSerThrThrThr 1820
Db 5402 GATCGAGAGCGCGACCGGATCGGAGCGGGAAGTCCATCTCTCATCGTCCACCGACG 5461
Qy 1821 ValGluHisAlaProIleTyrArgProGlyThrGluGlnSerSerGlySerSerGlySer 1840
Db 5462 GTGGAGCAGCGCCCATCTCGAGACTGTGTACAGAGCAGACGCGCAGCAGCGGAGC 5521
Qy 1841 SerGlyGlyGlyGlyGlySerSerSerArgProAlaSerHisSerHisAlaHisGlnHis 1860
Db 5522 AGCGGGCGGGGGGGGGGCGAGCAGCGCGCGCGCTCCCATCTCCCATGCGCCACGAC 5581
Qy 1861 SerProIleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeuHisAsn 1880
Db 5582 TCGCCCATCTCTCGCGACCCAGGATGCCCTCCAGCAGACGACCCAGTGTCTTCAAC 5641
Qy 1881 ThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeuArgSer 1900
Db 5642 ACAGGATGAGAGGTATCATCGCTGTGGAGCCAGCAGCCACCGTCTTGAGGTCC 5701
Qy 1901 ThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHisCysPro 1920
Db 5702 ACCTCCACCTCTCCACCGTTCGCCAGCTGCCACATTCCTCCACCTGCCACTGCCCA 5761
Qy 1921 LeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeuProLys 1940
Db 5762 CTGGCGGCGACCTCGATGGGGTCTACCTACCTCATGAGAGCGCTTGTGCTGCCAAG 5821
Qy 1941 GluAlaProArgValAlaAlaArgProGluArgProArgAlaAspThrGlyHisAlaPheLeu 1960
Db 5822 GAGGCCCCCGGTGCGCGCGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5881
Qy 1961 AlaLysProProAlaAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
Db 5882 GCCAAGCCCCCGCGCGCTCGGGCTGGAGCGCGCGCTCTCTCCCGCAGCAGGCGCTGGAG 5941
Qy 1981 ProArgProLeuValProProValSerGlyHisAlaThrIleAlaArgThrProAlaLys 2000
Db 5942 CCGCGCGCGCTTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6001
Qy 2001 AsnLeuAlaProHisHisAlaSerProArgProProAlaProProAlaSerAlaSerAsp 2020
Db 6002 AACCTCGACCTTACACCG 6061
Qy 2021 ProHisArgGlyLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSer 2040
Db 6062 CCGCACCGGGAAGAAGCTCAAAAGTAAACCTTTTTCATCCAGGAACCTGGAACTCCGTCT 6121
Qy 2041 LeuGlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSer 2060
Db 6122 CTGGGTACCAAGCAGCAGCTACAGCCCGGAGGGGTGGAGCGCGCTCAGCGCTGTGAGC 6181
Qy 2061 SerProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluLeuLeuAspLysSer 2080
Db 6182 TCACCCAGTCTTACCCACGACAGGGGCTCCCGCAGCAGCTTGGAGAGCTTGGAGAGG 6241
Qy 2081 HisLeuGluGlyLeuLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAla 2100
Db 6242 CACCTGAGGGGGAGCTGCGGCGCAAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 6301
Qy 2101 AlaHisLeuProHisLeuArgProLeuProGluSerGlnProSerSerSerProLeuLeu 2120
Db 6302 GCCACCTCCCGACCTTGGCGCGCGCTGCTGAGAGCAGCGCTGCTGCTGCTGCTGCTGCT 6361
Qy 2121 GlnThrAlaProGlyValLysGlyHisGlnArgValValThrLeuAlaGlnHisLysSer 2140
Db 6362 CAGACCGCGCGCGGGGTCAAAAGGTCAACAGCGGGTGGTCACTCCCTGGCGCGCACATCAGT 6421
Qy 2141 GluValIleThrGlnAspTyrThrArgHisHisProGlnGlnLeuSerAlaProLeuPro 2160
Db 6422 GAGGTCAATCACTAGGACTACACCGCGCACCCACAGCAGCTCAGCGCGCGCGCGCGCG 6481

SHSHAHQHSPISTQDALQORPSVLNHTQMKGIITAVEPSKPTVLRSTSTSSVPRPA
LEFAPATHCPLGTLGVYPTLMPEFVILLPKAEPRVARPERPRADTGHAFKAPRARS
LEPAPSSKGSSEPRLLVPVPSGHATIAITPAKNLAPHPASDPDPAPPASADPHREKT
QSKPFSIQELESLSLTHGSSYSPGVEPVSPSPSLTHDKLPGKLEBLDKSHLEG
ELRKPQPGVKLGGAHLPHLRPLPSQSPSPSLQTAQVKGQVRVTLAQHISEV
ITQDYTRHHPQOOLSAPLPAPLYSPFGASCPLVLRPPSDLYLPPDHGAPARGPHS
EGKRSPEPNKTSVLGGEDGIEPVSPGEMTBPGRSRAVTLXRDGEOTEPERSMG
SKSPNTSOPAPPSKLTESMSAMVSKKOEINKLANTHENEPEYNISOPGTELFNM
PALTCGLMTRYASOAVOEHASTNMGLEAIIJKALGKXDOHESPRISANAPENLAS
ASLPAAMPITADGSDHILSPGCGGKAKVSGRPSRKAASPAFGLASDRPPSVSS
VHSGDCNRRKPTLNRVWEDRPPSAGSTPPFPIINPLIMRLQAGVWASPPPPGLPAGSGP
LAGPHHAWDEBPFLKCSQYETLSDSE"

ORIGIN

Alignment Scores:

Pred. No.: 1,26e-254 Length: 8686
Score: 12978.00 Matches: 2487
Percent Similarity: 98.57% Conservative: 3
Best Local Similarity: 98.46% Mismatches: 9
Query Match: 98.21% Indels: 28
DB: 3 Gaps: 3

US-09-522-753-5 (1-2517) x AF125672 (1-8686)

Qy 1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
Db 157 ATGTGCGGCTCCACACAGCCTGTGGCACAGCTGTGGAGGCGCACTGAGCCCGCTACCCG 216
Qy 21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
Db 217 CCCACAGCCTTTCTTACCCAGTGCAGATCCCGCGAGCGCACAGGAGCTGGGCTCTG 276
Qy 41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
Db 277 GAGTACACAGCACCATCCCGGACATATGCTCCACACTGTCCCGCGGCTCCATCATCCAG 336
Qy 61 ProGlnArgArgProSerLeuLeuSerGluPheGlnProGlyAanGluArgSerGln 80
Db 337 CCCCAGCGCGGAGGCGCTCCCTGTCTGAGTTCGAGCGCGGGAATGAACGGTCCCGAG 396
Qy 81 GluLeuHisLeuArgProGlnSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
Db 397 GAGCTCCACCTGCGGCCAGAGTCCACATCACTGCTCCCGAGCTGGGAGTCAAGATG 456
Qy 101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
Db 457 GAGTTCATTTGAAGCAAGCGCCTCGGCTAGAGCTGTGCTGACCCCTCTCTCGGACCG 516
Qy 121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
Db 517 TCACCCCTGTGCGCACGCGGCGAGCTCGGGATCTGAAGACCTCACCAAGACCGTAGC 576
Qy 141 LeuThrGlyLysLeuGluProValSerProSerProSerProSerProSerProSerPro 160
Db 577 CTGAGCGGCAAGCTGGAGACCGGTGTCTCCCGGAGCCCGGACACTGACCTGAGCTG 636
Qy 161 GluLeuValProProArgLeuSerLysGluLeuIleGlnAanMetAspArgValAsp 180
Db 637 GAGCTGTGTGCGCCACCGGCTGTCAAGGAGGAGCTGATCCAGAAATGACGCGGTGAC 696
Qy 181 ArgGluIleThrMetValGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
Db 697 CGAGGATCACCATTGTAGGACGAGATCTTAAGCTGAGAGAGAGAGAGAGAGAGAGAG 756
Qy 201 GluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
Db 757 GAGGAGGAGGCTGCAAGCGCGCGGAGCTGAGAAGCGCGGTGTACCCCGCCCATCGAG 816
Qy 221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAanArgLysLysAlaGluAla 240
Db 817 TCGAAGCACCGCAGCCTGTGTGAGATCATCTACGACGAGAAACCGGAGAGAGGCTGAAGCT 876
Qy 241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260

Db 877 GCACATCGGATCTCGAAGAGCCTGGGCGCCCGAGGTGGAGCTGTCGTGTACACACGAGCC 936
Qy 261 SerAspThrArgGlnTyrHisGluAanIleLysIleAenGlnAlaMetArgLysLysLeu 280
Db 937 TCCGACACCCCGCAGTATCATGAGAACATCAAAATAAACACGAGCGATGCGGAGAGAGCTA 996
Qy 281 IleLeuTyrPheLysArgArgArgAenHisAlaArgLysGlnTyrLysGlnLysPheCysGln 300
Db 997 ATTCTGTACTTTGAGAGAGAGATCATCGCTCGAANAACANTGGAGAGAGTTCGTGCCAG 1056
Qy 301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAanLeuPro 320
Db 1057 CGCTATGACCACTCATGGAGGCGCTGGGAGAGAGAGGTGGAGCGCATCGAGAGAACACCC 1116
Qy 321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
Db 1117 CGCGCGCGCGCGAGGAGAGCAAGGTTCCGAGGTACTACGAGAGAGCAGTTCCTCTGAGATC 1176
Qy 341 ArgLysGlnArgGluLeuGlnArgMetGlnSerArgValGlyGlnArgGlySerGly 360
Db 1177 CGCAGCAGCGGAGCTGCGAGGAGCGCATGCGAG---AGGGTGGCGCCAGCGGCGAGTGGG 1233
Qy 361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleLeuAspGlyLeuSer 380
Db 1234 CTGTCCATGTCCTCCGCGCCGCGAGCAGCAGAGGTGTGAGAGATCATCGATGCGCTCTCA 1293
Qy 381 GluGlnGluAanLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
Db 1294 GACACAGAGAGCTGGAGAGAGCAGATGCGCCAGCTGGCGCTGTATCCCGCCCATCTGTGTAC 1353
Qy 401 AspAlaSerGlnGlnArgIleLysPheIleAanMetAanGlyLeuMetAlaAspProMet 420
Db 1354 GAGCTGACCAACAGCGCATCAAGTTTCATCAACATGAACCGGCTTATGCGCGACCCCATG 1413
Qy 421 LysValTyrLysAspArgGlnValMetAanMetTrpSerGluGlnGluLysGluThrPhe 440
Db 1414 AAGGTGTACAAATGACCCCGCAGGTCATGAACATGTGGAGTGGCGAGGAGAGAGACCTTC 1473
Qy 441 ArgGluLysPheMetGlnHisProLysAanPheGlyLeuIleAlaSerPheLeuGluArg 460
Db 1474 CGGAGAGAGTTCATGACAGCATCCCAAGACTTTGSCCTGTATCGCATCATCTCTCGAGAGG 1533
Qy 461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAanGluAanTyrLys 480
Db 1534 AAGACAGTGGCTGAGTGGCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1593
Qy 481 SerLeuValArgArgSerTyrArgArgGlyLysSerGlnGlnGlnGlnGlnGlnGln 500
Db 1594 AGCTTGTGTAGATCGAGCTATCGCGCGCGGCAAGAGCCAGCAGCAGCAACACAGCAGCAG 1653
Qy 501 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 520
Db 1654 CACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1713
Qy 521 AspGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGluLys 540
Db 1714 GATGAG 1773
Qy 541 AspLysGluAanLeuLeuLysGluLysThrAspAspThrSerGlyGluAanAspGlu 560
Db 1774 GACACAG 1833
Qy 561 LysGluAlaValAlaSerLysGlyArgLysThrAlaSerGlnGlyArgLysGly 580
Db 1834 AAG 1893
Qy 581 ArgIleThrArgSerMetAlaAanGluAlaAanSerGluLysLysLysLysLysLysLys 600
Db 1894 CGCATCACCCGCTCAATGGCTATGAGCGCAACAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 1953
Qy 601 SerAlaGluLeuAlaSerMetGluLeuAanGluSerSerArgTyrThrGluGluMet 620

Db 1954 AGCCCGAGCTGGCTCCATGGAGCTGAATGAGAGTTCTCGTGGACAGAGAGAAATG 2013
Qy 621 GluThrAlaLysGlyLeuLeuGluHisGlyArgAsnTyrSerAlaLeuArgMet 640
Db 2014 GAACAGCCCAAGAGGCTCTCTGGACACGGCCGCACTGGTGGCCATCGCCGGATG 2073
Qy 641 ValGlySerIleThrValSerGlnCysLysAsnPheTyrPheAsnTyrIleLysArgGln 660
Db 2074 GTGGGCTCCAGACTGTGTGCGAGTGAAGACTTCTACTTCACTACAGAGAGAGGCGAG 2133
Qy 661 AsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAlaArg 680
Db 2134 AACCTCGATGAGACTTTCAGACAGCAAGCTGAAGATGGAGAGAGAGCAACGCGCG 2193
Qy 681 ArgLysLysLysAlaProAlaAlaAlaSerGluGluAlaAlaPheProProValVal 700
Db 2194 AGGAGAGAGAGAGAGAGCGCGCGCGCGCGCGAGAGAGGCTGCAATTCGCGCCGCTGGTG 2253
Qy 701 GluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGluGlu 720
Db 2254 GAGGATGGAGATGGAGGCGTGGGCGTGCAGCGGAAATGAGAGAGAGATGGTGGAGAG 2313
Qy 721 AlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyProAla 740
Db 2314 GCTGAA-----GCC 2322
Qy 741 ThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaLys 760
Db 2323 ACTGTCAACACAGCTCAGACCGAGAGATCCCTCTCTCTCACTGAGGCGCCCAAG 2382
Qy 761 AspThrGlyGlnAsnGlyProLysProAlaThrLeuGlyAlaAspGlyProPro 780
Db 2383 GACACAGGCGCAATGGGCGCCCAAGCGCCAGCCAGCCCTGGGCGCGAGCGGCCACCCCA 2442
Qy 781 GlyProProThrProPro-ArgArgThrSerArgAlaProIleGlyProThrProAla 800
Db 2443 GGGCCACCCACCCACCCAGCGAGGACATCCCGGCC-CCCACTGAGTCCACCCCGGCTC 2501
Qy 800 rGluAlaThrGlyAlaProThrProProAlaProProSerProSerAlaProPro 820
Db 2502 TGAAGCCACTTAGCCCTTACGCCCTCCAGCGAGCCCTTCTCTCTCTCTCTCTCTCTCT 2561
Qy 820 oValValProLysGluLysGluGluThrAlaAlaAlaProProValGluGluGlu 840
Db 2562 TGTGTCTCCCAAGGAGAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2621
Qy 840 YGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGlu 860
Db 2622 GGAGGAGCAG 2681
Qy 860 uProValLysSerGlyCysThrGluGluGluGluGluGluGluGluGluGluGluGlu 880
Db 2682 GCCCGTCAAG 2741
Qy 880 aGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysGluGlyGly 900
Db 2742 GGAGGCGCTGAGGCGCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2801
Qy 900 rGlyArgAlaThrThrAlaLysSerGlyAlaProGlnAspSerAspSerAlaThr 920
Db 2802 CGGAGGCGCCACCAAG 2861
Qy 920 rCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuSerPr 940
Db 2862 CTGAGTGCAG 2921
Qy 940 oArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGlnLys 960
Db 2922 AAGGCCAGCCCTCTCACCGAGCTGGCGAGCCCGCGCGCGCGCGCGCGCGCGCGCG 2981
Qy 960 oLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaLeuProIleGlnValThrLys 980
Db 2982 ACTGGACCTGAAG 3041

Qy 980 sValHisGluLysProArgGluAspAlaAlaProThrLysProAlaProProAlaProPr 1000
Db 3042 AGTCCATGAGCTCCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3101
Qy 1000 oProProGlnAsnLeuGlnProGlnSerAspAlaProGlnGlnProGlnLysSerProAr 1020
Db 3102 GCCACCCCAAACTCTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3161
Qy 1020 gGlyLysSerAspSerProAlaProProAlaAspLysGlu----- 1033
Db 3162 GGGCAAG 3221
Qy 1034 ---AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProCysTyrThrSerG 1052
Db 3222 CCCAGCTTCCGCGCGAG 3281
Qy 1052 yLeuProPheProValProProArgGluValLysAlaSerProHisAlaProAspPr 1072
Db 3282 CTTGCCCTTCCGCGCGAG 3341
Qy 1072 oSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAspThrAl 1092
Db 3342 CTCAGCTTCTCTAGCTTCCACTGCTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3401
Qy 1092 aArgProValLeuProArgProProThrIleSerAsnProProProLeuLysSerSerAl 1112
Db 3402 CGGCGCGCTCTCCCGCGCCCAACCAACCTCTCAACCGCGCTCTCTCTCTCTCTCT 3461
Qy 1112 aLysHisProSerValLeuGluArgGlnIleGlyAlaLysSerGlnGlyMetSerValG 1132
Db 3462 CAAGCAGCCAG 3521
Qy 1132 nLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMetGlyLe 1152
Db 3522 GCTCCAGCTCTCTACTCAG 3581
Qy 1152 uProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGluLe 1172
Db 3582 GCGCTCTCCCTCTGAG 3641
Qy 1172 uSerProArgGlyGlnAlaGlyProProGlnSerLeuGlyValProThrAlaGlnGluAl 1192
Db 3642 GTCCCAAG 3701
Qy 1192 sSerValLeuAspGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGly 1212
Db 3702 GTCCGTCTGAG 3761
Qy 1212 eProSerThrAspValProSerAspSerAlaIleThrTyrArgGlySerIleThrHisG 1232
Db 3762 TCCAGCAG 3821
Qy 1232 yThrProAlaAspGluValLeuTyrLysGlyThrIleThrArgIleLysGlyGluAspSerPr 1252
Db 3822 CACGCCAGCTGAG 3881
Qy 1252 sSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluG 1272
Db 3882 GAGTCTGCTGGCGCGCGAG 3941
Qy 1272 yLysGlyGlyHisValLeuSerTyrGluGlyMetSerValThrGlnCysSerLysG 1292
Db 3942 CAAG 4001
Qy 1292 uAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgThrTyrAs 1312
Db 4002 GAGCGCAG 4061
Qy 1312 pMetMetGluGlyArgValGlyArgAlaLysSerSerAlaSerIleGluGlyLeuMetG 1332
Db 4062 CATGATGGAG 4121

QY 1332 YArgAlaIleProProGluArgHisSerProHisLeuLeuGluGlnHisHisIleAr 1352
 Db 4122 CGGTGCGATCCCGCGGAGGAGCACAGACCCCGCACCATCCAAAGAGCAGACCATCCG 4181
 QY 1352 gGlySerIleThrGlnGlyIleProArgSerTyValGluAlaGlnGluAspTyLeuAr 1372
 Db 4182 CGGCTCCATCAACAGAGGATCCCTCGGTCTTACGTGGAGGACAGAGGACTACCTCGG 4241
 QY 1372 gArgGluAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1392
 Db 4242 TCGGAGGCGCAAGCTCTTAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4301
 QY 1392 pLeuThrGluAlaTyIleThrGlnAlaLeuGlyProLeuLeuLeuLeuLeuLeuLeuLeu 1412
 Db 4302 CTGACCCGAGGCGCTTACAGAGCGAGGCGCTGGGCGGCGCTGAAGCTGAAGCGGCGGCTGA 4361
 QY 1412 uGlyLeuValAlaThrValIleGluAlaGlyArgSerIleHisLeuIleProArgGluG 1432
 Db 4362 GGGCTCTGTGGCCACGCTGAAGGAGGCGGCGGCTCCATCCATGAGATCCCGCGGAGGA 4421
 QY 1432 uLeuArgHisIleThrProGluLeuProLeuAlaProArgProLeuLeuGlySerIleTh 1452
 Db 4422 GCTGGCGACACGCGCGCTGCGGCGGCGGCGGCTCAAGAGGAGGCTCCATCAC 4481
 QY 1452 rGlnGlyThrProLeuLeuTyAspThrGlyAlaSerThrThrGlySerLeuLeuHisAs 1472
 Db 4482 GCAGGGCACCCCGCTCAAGTACGACACCGGCGGCTCCACCATGGCTCCAAAAGCACGA 4541
 QY 1472 pValArgSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspVa 1492
 Db 4542 CGTACGCTCCCTCATCGGAGCCCGCGGAGCGTTCACCCGCTGCACCGGCTGGATGT 4601
 QY 1492 lMetAlaAspAlaArgAlaLeuGluArgAlaCysTyIleGluLeuSerLeuLeuSerArgPr 1512
 Db 4602 GATGGCGACCGCGGCACTGGAGCGTCTGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGG 4661
 QY 1512 oGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleValProGl 1532
 Db 4662 AGGGACCCGACAGAGCTCGGGGGCTCCATTTGGCGGCGGCGGCGGCTCAATGTGCTCGA 4721
 QY 1532 uLeuGlyProArgGlnSerProLeuThrThrGluAspHisGlyAlaProPheAlaGl 1552
 Db 4722 GCTGGGTGAGCGGCGAGGCGGCTGACCTTATGAGGACACCGGCGGCGGCTTTGGCGG 4781
 QY 1552 yHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGlu 1572
 Db 4782 CCACCTCCACAGAGTTTCGCGCGTGACCATGGGAGGCGGCGGCGGCTGCGAGGAGG 4841
 QY 1572 ySerLeuSerSerSerLeuAlaSerGlnAspArgLeuLeuThrSerThrProArgGluI 1592
 Db 4842 CAGCCTTTCTGTCCAGCAGGATCCCGAGGACCGGAGGCGGCTGAGCTGCTGAGAT 4901
 QY 1592 eAlaLeuSerProHisSerThrValProGluHisPheHisProIleSerProTyrgl 1612
 Db 4902 CGCCAAAGTTCGCGGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4961
 QY 1612 uHisLeuLeuArgGlyValSerGlyValAspLeuTyArgSerHisIleProLeuAlaPh 1632
 Db 4962 GCACCTGCTTGGGGGCGTGAGTGGCGTGACCTGTATFCGAGGCGCACATCCCGCTGGCCTT 5021
 QY 1632 eAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaTyIleTyLeuPr 1652
 Db 5022 CGACCCCACTCCATACCCCGCGGCGATCCCTCTGACGCGAGCGGCTGCTACTACCTGCC 5081
 QY 1652 oArgHisLeuAlaProAsnProThrThrProHisLeuTyProProTyIleLeuArgGl 1672
 Db 5082 CGACACTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5141
 QY 1672 yTyProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyIleThrSe 1692
 Db 5142 CTACCCGACACGCGGCGGCTGGAGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 5201
 QY 1692 rGlnGlnMetHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGl 1712

Db 5202 GCAGCAGATGCCCAACACAGCGCACCGCATGGCGCCAGCGAGCTGATATGTGAGGGG 5261
 QY 1712 YLeuSerProArgGluSerSerLeuAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1732
 Db 5262 CCTCTGCGCGCGCGAGTCTCTGCTGGCATCAACTAGCTTGGGGTTCGCGAGGCGATCAT 5321
 QY 1732 eAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAl 1752
 Db 5322 CGACCTGTCTCCAGAGGCGCACCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5381
 QY 1752 aThrAlaMetAlaArgGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1772
 Db 5382 CACCGCATGGCGCGCTTGCCTACCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5441
 QY 1772 sSerSerSerProGlyGlyProThrHisLeuThrHisLeuThrHisLeuThrHisLeu 1792
 Db 5442 CAGCAGCTCCCGACTCTCCCGAGGAGTCCACACACACTTGCACAAAACACACACACGTC 5501
 QY 1792 rSerSerGluArgGluArgAspArgGluArgAspArgGluArgGluArgGluArgGluArg 1812
 Db 5502 CTCGTCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5561
 QY 1812 sSerIleLeuThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1832
 Db 5562 GTCCATCTCTCACTCCACACGCGGTGAGGACGACCCCATCTGGAGACCTGGTACAGA 5621
 QY 1832 uGlnSerSerGlySerSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1852
 Db 5622 GCAGACACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5681
 QY 1852 aSerHisSerHisAlaHisIleHisSerProIleSerProArgThrGlnAspAlaLeuGl 1872
 Db 5682 CTCCCATCTCCATGCGCACACGACTCGGCCCATCTCCCTCGGACCGGAGGATGCCCTCCA 5741
 QY 1872 nGlnArgProSerValLeuHisAsnThrGlyMetIleLeuLeuLeuLeuLeuLeuLeu 1892
 Db 5742 GCAGAGCCCATGTGCTTCAACACAGGATGAAGGATATCATCACGCTGTGGAGGC 5801
 QY 1892 oSerLeuProThrValLeuArgSerThrSerSerProValArgProAlaAlaTh 1912
 Db 5802 CACCAAGCCCGAGGCTCTGAGGCTCACCTCCACCTCTCCACCGGCTGCGGAGTGGCCAC 5861
 QY 1912 rPheProProAlaThrHisCysProLeuGlyIleThrLeuAspGlyValTyProThrIle 1932
 Db 5862 ATTCCACCTGCTACCCCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5921
 QY 1932 uMetGluProValLeuLeuProGlyAlaProArgValAlaArgProGluArgProAr 1952
 Db 5922 CATGGAGCGCGCTTGTGCTCCCAAGGAGGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 5981
 QY 1952 gAlaAspThrGlyHisAlaPheLeuAlaLeuProAlaArgSerGlyLeuGluProAl 1972
 Db 5982 AGGAGACCGGCTCATGCTCTCTCGCAAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6041
 QY 1972 aSerSerProSerGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
 Db 6042 CTCCTCCCGCAAGGGCTCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6101
 QY 1992 aThrIleAlaArgThrProAlaLeuAsnLeuAlaProHisAlaSerProAspProPr 2012
 Db 6102 CACCATCGCGCGACCCCTCGGAAGAACCTCGACCTCACACGCGGCGGCGGCGGCGGCGG 6161
 QY 2012 oAlaProProAlaSerAlaSerAspProHisArgGluLeuThrGlnSerIleProPheSe 2032
 Db 6162 GGGCGACCTGCTCGGCTCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6221
 QY 2032 rIleGlnGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 2052
 Db 6222 CATCAGGAGCTGAACTCCGCTCTCTGCGGTACCGCGGAGGAGGATACAGCCCGGAGG 6281
 QY 2052 YValGluProValSerSerProSerLeuThrHisAspLeuGlyLeuProTy 2072

Db	6282	GGTGGAGCCGCTGAGCTCAGCCAGTCTGACCCAGTCTGACCCAGCAGCAGGAGGCTCCCCAA	6341
Qy	2072	sHleLeuGluLeuAepLySerHisLeuGluGlyLeuLeuAepProLySerGluProgl	2092
Db	6342	GCACCTGGAGAGCTCGACAGAGCCACCTGGAGGAGGAGCTGGCGGCCAGCAGCCAGG	6401
Qy	2092	yProValLyLeuGlyGlyGluAlaHisLeuProHisLeuAepProLeuProgluSe	2112
Db	6402	CCCGGTGAAGCTGGCGGAGGAGCGCCACCTCCACACCTGGCGGCGCTGCCTGAGG	6461
Qy	2112	xGlnProSerSerProLeuLeuGlnThrAlaProGlyValLyHisLeuAepGva	2132
Db	6462	CCAGCCCTGCTCCAGCCGCTGCTCCAGAGCCGCGCCAGGGGTCAAGGTCACCCAGGGGT	6521
Qy	2132	lValThrLeuAlaGlnHisLeuSerGluValLeuThrGlnAepLyThrArgHisLeuPr	2152
Db	6522	GGTACCCCTGGCGCCAGCACATCAGTGGGTGATCATCACAGGACTACACCCGAGCACCC	6581
Qy	2152	oGlnGlnLeuSerAlaProLeuProAlaProLeuLySerPheProGlyAlaSerCyePr	2172
Db	6582	ACAGCAGCTCAGCGCACCCCTGCGCGCCGCTCTACTCTCTCCCTGGGCGCAGCTGCC	6641
Qy	2172	oValLeuAepLeuAepArgProProSerAepLeuLyLeuProProAepHisGlyAl	2192
Db	6642	GTCTCTGGACCTCCCGCCGCCACCCAGTGACCTCTACTCTCCCGCCCGCCGAGCCATGTTGC	6701
Qy	2192	aProAlaArgGlySerProHisSerGluGlyGlyLyArgSerProGluProAepHisGly	2212
Db	6702	CCCGGCCGCTGCTCCCGCCGCCACCCAGTGAGGAGGAGGCTCTCCAGAGCCAAACAGAC	6761
Qy	2212	rSerValLeuGlyGlyGlyGluAepGlyGlyGluProValSerProProGlyGlyMetTh	2232
Db	6762	GTCTCTGCTGGGTGGTGGTGGAGGAGGAGTATGAGCTGTGTCTCCCGCCGAGGAGCATGAC	6821
Qy	2232	xGluProGlyHisSerArgSerAlaValTyProLeuLeuTyArgAepGlyGluGlnTh	2252
Db	6822	GGAGCCAGGAGGAGCTCCCGAGTGTGTGTACCGCTGTGTACCGAGTGGGAGACAGAC	6881
Qy	2252	xGluProSerArgMetGlySerLySerProGlyAepThrSerGlnProProAlaPhePh	2272
Db	6882	GGAGCCAGCAGGATGGCTCCAAAGTCTCCAGGACCAACAGCAGCCAGCGGAGCTCTT	6941
Qy	2272	eSerLyLeuThrGluSerAepSerAlaMetValLySerLyGlyGlnGlyLeuAepLy	2292
Db	6942	CAGCAGCTGACCGAGAGCACTCCGCGCATGCTCAAGTCCAGGAGGAGGAGGATCAACA	7001
Qy	2292	sLyLeuAepThrHisAepAepGluProGluTyArgHisLeuSerGlnProGlyThCl	2312
Db	7002	GAGCTGAGACCCACACCCGAGTGGCTGATACATATACAGCCGCTGGGAGCGGA	7061
Qy	2312	uilePheAepMetProAlaLeuThrGlyLeuMetThrTyArgSerGlnAlaVa	2332
Db	7062	GATCTTCAATATGCGCGCATCAACCGGAAACAGGCTTATGACCTATAGAGCAGCGCGT	7121
Qy	2332	lGlnGluHisLeuAepSerAepMetGlyLeuGluAlaLeuAepGlyAlaLeuMetGly	2352
Db	7122	GCAGGACATGACGACACACATGCGGCTGGGAGGAGGATATTAGAGGAGCTCATGG	7181
Qy	2352	yLyTyAepGlnThrGluGluSerProProLeuSerAlaLeuAlaPheAepProLeuAs	2372
Db	7182	TAAATATGACAGTGGGAGAGTCCCGCGCTGAGCCGAGCCATCTCTTTAACCTCTGAA	7241
Qy	2372	nAlaSerAlaSerLeuProAlaAlaMetProLeuThrAlaAlaAepGlyArgSerAepHi	2392
Db	7242	TGCCAGTGGCCGCTGCGCTGCTATGCGCATACCGCTGCTGACGAGCGAGTACCA	7301
Qy	2392	sThrLeuThrSerProGlyGlyGlyGlyAlaValSerGlyArgProSerSerAr	2412
Db	7302	CACACTACTCTGCGAGTGGCGGAGGAGGAGGAGGAGTCTCTGCGAGCCAGCGAGCG	7361
Qy	2412	gLyAlaLySerProAlaProGlyLeuAlaSerGlyAepArgProProSerValSerSe	2432
Db	7362	AAAGCCAGTCCCGCGCCCGCGCTGGCATCTGGGAGCGGCGCCCTCTGTCTCTCTC	7421
Qy	2432	rValHisSerGlyGlyAepCyAsnAepAepGlyThrProLeuThrAsnAepValTtpGluAs	2452
Db	7422	AGTGACCTCGGAGGAGGAGCTGCACACCGCGGAGCGCTCCACACCGCTGTGGAGGA	7481
Qy	2452	pArgProSerSerAlaGlySerThrPheProTyArgProLeuProLeuMetArgLeuGl	2472
Db	7482	CAGGCCCTCTGTCGAGGTTCCAGCCATTCCTCCATCAACCCCTGATCATGCGGCTGCA	7541
Qy	2472	nAlaGlyValMetAlaSerProProProGlyLeuProAlaGlySerGlyProLeuAl	2492
Db	7542	GGCGGCTGTCTGCTTCCACACCCCGGCTCCCGCGGCGGAGCGGCGCTCTGC	7601
Qy	2492	agLyProHisLeuAlaTtpAepGluGluProLySerProLeuLeuCySerGlnTyGluTh	2512
Db	7602	TGGCCCGCCAGCTGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	7661
Qy	2512	rLeuSerAepSerGlu 2517	
Db	7662	ACTCTCGACAGGAG 7677	
RESULT 3			
LOCUS	AF113001	8388 bp mRNA linear ROD 20-MAR-1999	
DEFINITION	Mus musculus silencing mediator of retinoic acid and thyroid hormone receptor alpha mRNA, complete cds.		
ACCESSION	AF113001		
VERSION	AF113001.1	GM 4454547	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 8388)		
AUTHORS	Ordentlich, P., Downes, M., Xie, W., Genin, A., Spinner, N.B. and Evans, R.M.		
TITLE	Unique forms of human and mouse nuclear receptor corepressor SMRT		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)		
MEDLINE	99178941		
PUBMED	10077563		
REFERENCE	2 (bases 1 to 8388)		
AUTHORS	Downes, M.R., Ordentlich, P. and Evans, R.M.		
TITLE	Direct Substitution		
JOURNAL	Submitted (11 DEC-1998) Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES	Location/Qualifiers		
source	1..8388		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
	/tissue_type="spleen; brain"		
	626..18047		
	/function="transcriptional co-repressor"		
	/note="mSMRT alpha"		
	/codon_start=1		
	/product="silencing mediator of retinoic acid and thyroid hormone receptor alpha"		
	/protein_id="AAD20944.1"		
	/db_xref="GI:4454548"		
CDS	/translation="MSGTQFVACTWRAAPRYPHGISYVQIARSHDVGLELVEQVH HPDMVSHLSPGSIQIPRRPRLSEFPQSGSEUHLHPESRFLPSLGPDI TESKPRLEPLDPLRPSLPATPGSGEDLTQDRLAGLEPVPFPSPHAPDEL ELAPSLKEELIQMDRVREITWVEQI SKLKKQQQLEEAAPKEPVPSP IESKLSLVQIYIDNRKAAAHLEGLGPVELPLYNQSDTQYHENIKINAM RKCLLYFKRRHARKMORFQRYDQLEMAWEKVEIENNPRAKESKVEYTE KQPEFKRLOERMSRVQSGLSMSAARSHVESEIIDGLSEQENLEKQKRL AVIPNLYDADQRIKTINMGLDDPMKYVDQVTNMSEQRDTFRKFWHPKN FGLIPLFLPKTKVACVLYLYTKQENYKSLVRRYRRGKSGSQQQQQQQQQ ARSGEKE TANSGRRKRLYSRMANEHNSETATPOSSSELSMEMNESSEMTSEMETAKKGL EHRNWSAIRHVRGKTSCKQKPYFYNYKRNQLDELIOQHLKQEVNARKKKKT PAAASEETAPPADEMEASGASNEBELAEASQASQSGNEVFRVGCSPAAVN		

1217 lProSerAspSerAlaIleThrTyrArgGlySerIleThrHisGlyThrProAlaAspVa 1237
1492 TGACAGCGGCCCC-----AGCTACAGAGGCTCTATACCCACGGCAGCGCCGACAGCT 4245
1237 lLeuTyrLysGlyThrIleThrArgIleIleGlyGluAspSerProSerArgLeuAspAr 1257
4246 CCTCTACAAGGTTACCATCAGCAGGATCGTGGTGAGACAGCCCAAGTCGCTTGACG 4305
1257 gGlyArgGluAspSerLeuProLysGlyHisValIleTyrGluGlyLysGlyHisVa 1277
4306 GCACAGAGAGACACCTGCCCAAGGGCCATGTCTATATAGGGGCAAGAAAGGCCACGT 4365
1277 lLeuSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSe 1297
4366 CCTATCCTTATGAAGTGTATGTCTCGTGTCTACAGTGTCTTAAGGAGGATGCAAGGACG 4425
1297 rSerGlyProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyAr 1317
4426 CTCGGGCCACCATCAGACTGCGGCCCTTAACGCACCTATGACATGATGGAGGCCG 4485
1317 gValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProPr 1337
4486 TGATAGGAGGACTGTCTACCTCAGCCAGCATAGAGGACTCATGGGCCGCGCATC---CC 4542
1337 oGluArgHisSerProHisGlyLeuLysGluGlnHisIleIleArgGlySerIleThrGl 1357
4543 TGAGCAGCACAGCCCC---CACCTCAAGGAGCAGCATCATCCGAGGCTTCATCAGCA 4599
1357 nGlyIleProAspSerTyrValGluAlaGlnGluAspTyrLeuArgArgGluAlaLysLe 1377
4600 AGGCATCCCGAGGTCTTATGTGAGGCGCAGAGACTTACTACGGCGGAGGCGCAAGCT 4659
1377 uLeuLysArgGluGlyThrProProProProProProSerArgLeuThrGluAlaTy 1397
4660 CTTGAGCGGAGAGGAGACACACCCACCCACACACCTCGGGGACCTGATGAGACCTA 4719
1397 rLysThrGln-----AlaLeuGlyProLeuLysLeuLysProAlaHisGluGlyLe 1414
4720 CAAGCCCGCGCCCTCGACCTCTGGGTCCCTCGAAGCTGAAGCCGACTCACGAGGCTGT 4779
1414 uValAlaThrValIleGluAlaGlyValGlySerIleHisGluIleProArgGluLeuAr 1434
4780 GTTAGCACTGTGAGAGGCGCGCGCTCTATCCATGAGATCCCGAGAGGAGCTGCG 4839
1434 gHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGl 1454
4840 CGCACACCTGAGCTACCTCGGCACCAACCGCTCTGAAGGAGGGTTCCATCACCAGGG 4899
1454 yThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValAr 1474
4900 CACCCCACTCAAGTACGACTCTGGGGCACCTCCACTGGCACCAGAGAAACACGACTGG 4959
1474 gSerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAl 1494
4960 CTCATCATCGGAGCGCGCGCGCTTTCCCTGCGCCCTGACCCCGCTGGACATAATGCG 5019
1494 aAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSerArgProGlyTh 1514
5020 TGACGCGCGGGCCTGGAGCGTGCCTGCTATGAAGAGAGTCTGAAGAGCGGCTCAGGGAC 5079
1514 rAlaSerSerSerGlySerIleAlaArgGlyAlaProValIleValProGluLeuGl 1534
5080 CAGCAGTGTGCGAGGGGGTCCATCAACGTGGGGTCCAGTGTGCTGCTGACTGAGTGGG 5139
1534 yLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPheAlaGlyHisLe 1554
5140 CAAGCCACGCAAGGCCACTGACTTACGAAGACCAACCGGGGACCCCTTCACCAAGTCACT 5199
1554 uProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLe 1574
5200 GCCACGTGTCTCCCTGTGACCAAGGAGGAGGCCACCGCCAGCGCTTCAGGAAGGACGCT 5259

1574 userSerSerLysAspSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLy 1594
5260 CCTATCCAGCAAGCGCTCCAGGACCGGAAGCTGACATCTACACCCCGGAGATCGCCAA 5319
1594 sSerProHisSerThrValProGluHisHisProHisProHisSerProTyrGluHisLe 1614
5320 GTCCCAACACAGGACTGTGCGCCGAGCAACACCTTACCCCATCTCCCCCTATGAGACTT 5379
1614 uLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLeuAlaPheAspPr 1634
5380 GCTCCGGGGCGGACTGTGTGTGACCTGTACCGTGTGCATCCCATTTGGCTTTGACCC 5439
1634 oThrSerIleProArgGlyIleProLeuAsp-----AlaAlaAlaAlaTyrTyrLeuProAr 1653
5440 CACTTCCATACCTCGAGGGATCCTCTGGAGACGAGCCGAGCTTACTTCTGCCCCG 5499
1653 gHisLeuAlaProAspProThrTyrProHisIleuTyrProProTyrLeuIleArgGlyTy 1673
5500 GCATTTGGCCCGGAGGCCACCTTACCCACACCTGTACCCACCTTACTCATCTCGGGCTA 5559
1673 rProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerGl 1693
5560 CCTGACACGGGGCCCTTGGAGAACCGCCAGACCATCATCAATGACTTACATCACCTGCA 5619
1693 nGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLe 1713
5620 GCAGATGCACCCGCAACCGTCCCTCCGCCATGCGCCAGCGTGTGACATGCTGAGGGGTCT 5679
1713 userProArgGlnHisSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleAs 1733
5680 GTACCGCGAGAGTCTCGCTGGCCCTCAATATATGCCGCTGGCCCAAGAGGATATATGCA 5739
1733 pLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaTh 1753
5740 CCGTGTCCCAAGTGCACACCTGCGCCGTGTGTGTGCCAACCAACGACGAGCACCTGTCAC 5799
1753 rAlaMetAspAlaLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisLe 1773
5800 CGCCATCGACCGCTTGGCTTACTTCCCACTCGCCCGCCACCTTTCAGCAGCGCCGACAG 5859
1773 rSerSerProHisSerProGlyGlyProThrHisIleuThrLysProThrThrThrSerSe 1793
5860 TAGCTACCGCTGTGCGGAGGAGGCCCTCCTACCTAGCTAAACCACTGCCACATCTTC 5919
1793 rSerGluArgGlnArgAspArgAspArgGluArgAspArgGluArgGluLysLe 1813
5920 ATCGGAGCGGAGCGGAACTGAGCGGGAACGAGAC-----AAGTC 5961
1813 rIleLeuThrSerThrThrValGluHisAlaProIleTyrArgProGlyThrGluGl 1833
5962 CATCTCACCTTACCACTACAGTGTGAGCATGACCCCATCTGGAGACTGTGTACGAGCA 6021
1833 nSerSerGlySerSerGlySerSerGlyGlyGlyGlySerSerSerSerArgProAlaSe 1853
6022 GAGCAGCGGGGCT-----GGGGGCGCAGCGCCCGCCGCTC 6057
1853 rHisSerHisHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnGl 1873
6058 CCACAC-----CACACACCTTCCCTCCCGGAGCCAGGACGCTTTCAGCA 6111
1873 nArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSe 1893
6112 GAGGCCCGGAGTGTCTGCACACAGCATGAGGGGTGGTGCATCTCCGTGGAAACCCGG 6171
1893 rLysProThrValLeu-----ArgSerThrSerThrSerSerProValArgProAl 1910
6172 CACGCCACCGGCTCTGAGGTGGGCCAGGCTCCACCTTCCACCTTCTCGCTTCCGCCGACG 6231
1910 aAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrPr 1930
6232 TGCCACATTCCTCCCTGCCACCACTGCCCCACTTGTGTGACCCCTTGAAGGGGTCTACCC 6291
1930 oThrLeuMetGlnProValLeuLeuProLysGluAlaProArgValAlaArgProGluAr 1950

Db 6292 TACCCTCATGGAGCGCTCTGTTACCAAGGAGAGCTCTGGGTGCGCGCGCGAGCG 6351
 Qy 1950 qProAaGlaAaPThrGlyHisAlaPheLeuAlaLysProProAlaAaSerGlyLeuG1 1970
 Db 6352 GGCCCGGGTGGAGCTGGCCATGCTCTTACCAAAACCCCGGGCGG-----GA 6402
 Qy 1970 uProAlaSerSerProSerLysGlySerGluProAaGProLeuValProProValSerG1 1990
 Db 6403 GCGCGCTCTTACCCAGCAAGAGCTCCGAGCGCGCGATCCCTTAGCACCCCGAGCTCCAG 6462
 Qy 1990 yHisAlaThrIleAlaAaGThrProAlaLysAlaLeuAlaProHisHisAlaSerProAs 2010
 Db 6463 CCACACAGCCATCGCGCGCAGCCCGCAAGAAAGACCTTGCACCCACCCATCCAGTCCGGA 6522
 Qy 2010 pProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysPr 2030
 Db 6523 CCGCGCGCGG---CCACCTCGGCTCAGATCTGCACCGAGAAAGACTCAAGTAAC 6579
 Qy 2030 oPheSerIleGlnGluLeuAlaArgSerLeuGlyTyrHis---GlySerSerTyrSe 2049
 Db 6580 CTTTTCATCCAGGAATTGGAACTCCGTCTCTGGGTACACAGAGTGAGCTGGCTACAG 6639
 Qy 2049 rProGluGlyValGluProValSerProValSerProSerLeuThrHisAlaPlyeG1 2069
 Db 6640 CCGCATGGGTGGAGCCCATCAGCCCGGTGAGCTCCCGCAGCTGACCCACGACAGAGG 6699
 Qy 2069 yLeuProLysHisLeuGluGluLeuAlaPlySerHisLeuGluGlyLeuArgProLy 2089
 Db 6700 GCTCTCCAACTCTGGAGAGCTAGAGAAAGCCACTTGGAGGGAGCTGGCGACAA 6759
 Qy 2089 sGlnProGlyProValLysLeuGlyGlyGluAlaHisLeuProHisLeuArgProLe 2109
 Db 6760 GCAGCAGGCGCCCAAGAGCTCAGCGGGAGGTGGCCATCTCCACATCTGGCGCACT 6819
 Qy 2109 uProGluSerGlnProSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHI 2129
 Db 6820 GCCGAGGCGCAGCCCTCCTCCAGCCCACTCTCCAGAGTCCCGCAGGCATCAAGGTCA 6879
 Qy 2129 sGlnArgValThrLeuAlaGlnHisIleSerGlnValIleThrGlnAspTyrThrAr 2149
 Db 6880 CCAGAGGTGGTCACTGGCTCAGACATCAGCAGAGTCAATAGCAGAGCTACACGG 6939
 Qy 2149 gHisAlaProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAl 2169
 Db 6940 CCACCAACCGCAGAGCTCAGTGGCGCCCTTCCCGCCCTCTCTACTCTCTTCCCGGAGC 6999
 Qy 2169 aSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProAs 2189
 Db 7000 CAGTGGCTGCTGCTGGATCTTGGCGCCCAACCCAGTGACCTTACTCTCCACCCCGCA 7059
 Qy 2189 pHisGlyAlaProAlaArgGlySerProHisSerGluGlyLysArgSerProGluPr 2209
 Db 7060 CCATGGCACCACCGCGCGGATCCCCCAGTGAAGGGGGCAAAAGGTCCCGAGAAC 7119
 Qy 2209 oAsnLysThrSerValLeuGlyGlyGluAspGlyIleGluProValSerProProG1 2229
 Db 7120 CAGCAAAACATCGGTCTGGGCGAGCGAGGATGCCATTGAGCTGTGTGCCACCAAGA 7179
 Qy 2229 uGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspG1 2249
 Db 7180 GGGATGACTGAGCAGGACATGCTCGAGCACTGGTACCACTGCTGTATGAGACGG 7239
 Qy 2249 yGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProPr 2269
 Db 7240 GGAACAGGCGAGCC---AGGATGGGTCTAGAGTCTCCAGGCAACACCCAGCGCGCC 7296
 Qy 2269 oAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnG1 2289
 Db 7297 AACCTTCTCAGTAGCTGACTGAGAGCACTCCGCGCATGTGTGAGTGAAGAGCAGA 7356
 Qy 2289 uIleAsnLysLeuLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnPr 2309

Db 7357 GATCAACAAGATCTCAACACCCAGCAAGCGAGCCAGCAATACAATATGGCCAGCC 7416
 Qy 2309 oGlyThrGluLysPheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSe 2329
 Db 7417 TGGGACGGAAATTTTCAACATCGCGCATCTCTGGAGCAGCGCTTATGACCTGTAGAAG 7476
 Qy 2329 rGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleLeuArgLysAl 2349
 Db 7477 CAGCGGTGCTGGAACCGCGACCAACATGGGGCTAGAGGCATATTAGAAGCC 7536
 Qy 2349 aLeuMetGlyLysTyrAspGlnTTPGluGluSerProProLeuSerAlaAsnAlaPheAs 2369
 Db 7537 ACTCATGGGTATATGATCATGGGAAGAGCCCGCGCTCGCGCGCAATGCTTTTAA 7596
 Qy 2369 nProLeuAsnAlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspG1 2388
 Db 7597 CCTCTGAATGCGAGCGCCAGTCTGCCCGCTGCTGTATGCCATAACCACTGCTGACGG 7656
 Qy 2388 yArgSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyAr 2408
 Db 7657 ACGAGTGACCACTGCACTCCTCGCAGGTGGAGGAAAGCAAGGTCTCTGCAG 7716
 Qy 2408 gProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProPr 2428
 Db 7717 ACCTAGCAGCCCAAGCCNAGTCGCCAGCACCCAGGCTAGGCTCGGAGACCGACCC 7776
 Qy 2428 oSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnAr 2448
 Db 7777 TCTGTCTCTGCTGACTCAGAGGGGAGCTGCATTCGCCGACACCACTACACACCG 7836
 Qy 2448 gValTTPGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIl 2468
 Db 7837 TGTGTGGAGGTCGCGCCCTCATCTGCAGGCTCACGCCATTCCCTACAAACCTTTGAT 7896
 Qy 2468 eMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySe 2488
 Db 7897 TATGAGGTACTAGGAGGTGTCTGGCTTCCCGCCCGCCCACTGGCTTGGCGAGCGAG 7956
 Qy 2488 rGlyProLeuAlaGlyProHisAlaTTPAspGluGluProLysProLeuLeuCysSe 2508
 Db 7957 CGGCGCCCTAGCTGGTCCCACCGCTGGATGAGAGCCCAAGCCACTGCTGTGTTTC 8016
 Qy 2508 rGlnTyrGluThrLeuSerAspSerGlu 2517
 Db 8017 ACAGTATGAGAGCTCTCGACAGCGAG 8044

RESULT 4
 AF125671 8544 bp mRNA linear ROD 04-APR-1999
 LOCUS Mus musculus silencing mediator of retinoic acid and thyroid
 DEFINITION hormone receptor extended isoform (Smrte) mRNA, complete cds.
 ACCSSION AF125671
 VERSION AF125671.1 GI:4559295
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 8544)
 Park, E. J., Schreen, D. J., Yang, M., Li, H., Li, L. and Chen, J. D.
 SMRte, a silencing mediator for retinoid and thyroid hormone
 receptors-extended isoform that is more related to the nuclear
 receptor corepressor
 Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)
 MEDLINE 99199215
 PUBMED 10097068
 REFERENCE 2 (bases 1 to 8544)
 Chen, J. D.
 Direct Submission
 TITLE Submitted (03 FEB-1999) Pharmacology and Molecular Toxicology,
 University of Massachusetts Medical School, 55 Lake Avenue North,
 Worcester, MA 01655, USA
 JOURNAL Location/Qualifiers
 FEATURES

source

1. 8544
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1. 8544
/gene="Smrte"
160..7548
/function="transcriptional corepressor"
/note="SMRte"
/codon_start=1
/product="silencing mediator of retinoic acid and thyroid hormone receptor extended isoform"
/protein_id="AAD22972.1"
/db_xref="GI:4559296"

gene

CDS

HPDYTHLSPGSI IQPDRRPSLLSFPQPSERSDELHLPSERTLPELGLKDELP
TESRPELELLPTLRLSPILLATGQPSGSEDLTKDRSLACKLSPVSPSPHAPDEL
ELASRLSKELIQNLRDLREBEITVYQOISKLKKQOOLEEBAAPPEPKSPVP
PIEKHSLVQLIYDENRKAARAHRIEGLGPQVPELPIQPSQTKQHENIKNOA
MRKLILYFKRNHARKQWQRCQYDQJLMEAWKKVERIENPRRRKESKVRREY
EKQPEIRKQRELOERMOSRGSGLSMAARSSEHSIIDLSEBQENLEKQHRQ
LAVIRHVRDQORIKFINMGLMDPMKYDQRTVMWSEBQERTFREKFWOHPKN
FGLTASFLERKTVAECLVYLYTKKNENKSLVRSYRGRKSGQOQOQOQOQOQ
ARSSQBEKE
TANSQGRKGRITRSMANEANHEATATPOOSBELAMENNESRWTEEMETAKGLL
EHNENWAIARMYKSTVSOCKNFVNYKRONLDEILQKHLMKERNARKKKV
PAASERTAPPAADEMEASGASANEELAEASQASGNEVPVGRCSGPAVN
NSDTEPSVPSSEATKQPKPTGTEALPAATQPPVPPREPAAPAEPSVPAASG
PPSPPEPSAPPAATVDKQEPAPAPAPQTEDAKESAEEDIDVKEPEPEASEE
PPSVKSDHKEETEEPEKNGRLELTETVSEAPLKYERAGSKAATNGSGSSTQDS
DSSATCADEVEDEGGDGRLLSPSLITPAGDPRASTSPQKPLDLKQKRAAI
PPVITVKEPREDTVPKPPVPPPTPQLOPBGDVSGQSGSPKSGSPSPVPAEK
BAEPAPFAPFTEGQSLPHAGHELPSHPREVIKTSTRADPLFSYTPPGHPLP
GLHDSAPVLPRIINPNPPLISSAKHPGLERLGAISQMSVQLRVPHSHEAKPMG
LIMELPLAVDPKKLTALAPPVVEASPRASQVPCRRPQLRLYHPRTPADVLVYGT
ISRTVSDSPRLDRAREDTLPKGVITYEGKGHVLVSVEGMSVSCSKEDGRSSGP
PHEAAPQRYDMMGRVGRVTSASIEGLMGRAIPBOHSPLKXOHHIRGSTTGIP
RSVYAEQDYLREAKLLKRGSTPPPPPPRDLTETYPKPLDPLGLKLPKTHGV
ATVKEAGSIHEIPRELRTEPLAPRLPKEGSIOTGTPKLYDSCAPSTGKHDV
RSIIGSGRFPFALHPLDIMALALACERACESLSKSGTSAGSGSITRGAPVVP
ELGPROSLTYEDGAPTHLSPGSPVITREPTPQOGLSLSSKASQDKLSTP
REIASKPSHTVPEHPHPISTPEHLGRVTVGLYHGHIPLAFDPTISIPGILBAA.
AAYDLPLHAPSPTPHLYPYLYGYDPTAALNNQTIINDYITSQOMHNAAMA
QADMLGSLPRESLALNYSAGPRGIDLSQVHLPLVLPPTPTGTPATADILHAP
TAPPPSSHSGLSPGCPHAKPTATSSERERERERERERERERERERERERER
WRPTEQSGAGSGSRPASHTHQSPISPRTDALQQRPSVLNHTSMKGVTSVFGT
PTVLRSSTSSPVPAATFPFATHCPGLGTEGVFTLMEPVLPLKETSVPARPERP
VDGHAFLTKPPARPSPSGSSEPSLAPSSHTLARTPAKSLAPHASPPDPG
PTASDLHREKTKOSKPFQIOLSLRSLGHSAGYSPOGVRPISPVSPSLTHDKGLS
KPLBELSKHLELRHQPQPKMLSAEHLPLRLPLPESQPSSSLQOTAPGLKH
QVUTLQAHLSEVITQDTRHHPQOLSGPLPABLSFGASCPVLDLRPRSDVLP
PDHGTARGSPHSBEGKSPSPKTSVLGSSDEAIEPVSPEGMEFCHARSVPPLL
YRQGEQSPMGSGFNSTQSPAFKLTSSNASKVSKQKQKGNKGNKGNKGNKGN
NIGQFTEIFNMASITAGMLTCKSQVQAEHASTNGLEAIRKMGKDYDQNEPPP
LGNAPFLNMAATPAAPAAEMITADGSDHASTLTPGGGKAKVSGSPSRKAKSP
GLASGDRPSSVSHSGEDCNRTPLNRYWEDRPSAGSTPPFPYPLNRLQAGVMA
SPPPPGLAAGSGPLAGPHAWDEBPKEKLLCSQVETLSDSE"

ORIGIN

Alignment Scores:

Pred. No.: 3,11e-211 Length: 8544
Score: 10822.50 Matches: 2149
Percent Similarity: 87.94% Conservative: 90
Best Local Similarity: 84.81% Mismatches: 197
Query Match: 81.97% Indels: 117
DB: 10 Gaps: 33

US-09-522-753-5 (1-2517) x AF125671 (1-8544)

Qy 1 MetSerGlySerThrGluLeuValAlaGlnThrTyrArgAlaThrGluProArgTyrPro 20
Db 160 ATGTCCAGGATCCACACAGCTGTGGCACAGACATGGCGGCTGCTGAGCCCCCGCTACCCA 219

Qy 21 ProHisSerLeuSerTyrProValGlnIleAlaAaThrHisThrAspValGlyLeuLeu 40
Db 220 CCCCATGGCATCTCTTACCCTGGGTGAGATAGCCCGGTCCACACGAGCGTGGCGCTGCTT 279
Qy 41 GluTyrGlnHisHisSerArgArgTyrAlaSerHisLeuSerProGlySerIleLeuGln 60
Db 280 GAGTACCAACACACACCCCGGTACTAACCTCACACCTGTCAACCGGTTCATCATCCAG 339
Qy 61 ProGlnArgArgTyrProSerLeuLeuSerGluPheGlnProGlyAengluArgSerGln 80
Db 340 CCACAGAGGAGTGGCCCTCACTGCTGTCCAGAGTTCACAGCTGGGAGTGAAGCGTTCAG 399
Qy 81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
Db 400 GAGCTCCACCTGGCCCTGAGTCCCGCAGCTTCTGCTGAGCTGGGCAAGCCGACATA 459
Qy 101 GluPheIleGluSerLysArgArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
Db 460 GAATTCACCGAGCGCAAGCGCCCGCTGGAGCTACTACCCGATACCTCTGCTGGGCCCA 519
Qy 121 SerProLeuLeuIleThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
Db 520 TCACCCCTGCTGCCACTGGGCGACGAGTGGGTCTGAAGACCTTACCAAGGACCGGTAGC 579
Qy 141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
Db 580 CTGGCAGGCACTGGGAGCTGTGTCACTCCCTGCTGAGCTGGGCAAGCCGCTGAGCTA 639
Qy 161 GluLeuValProArgLeuSerLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 179
Db 640 GAGCTGGCCGCTCTGACCTGTCCAGGAGGAGCTGATCCAGAACAGATGACCGCGCTG 699
Qy 180 AspArgGluLeuThrMetValGluGlnGlnIleSerLysLeuLysLysGlnGlnGln 199
Db 700 GACCGTGTAGATGCATCATGTGTAGACAGCAGCAGATCTCCAAGCTGAAGAGAGAGCAACAG 759
Qy 200 LeuGluGluGluAlaAlaLysProGluProGluLysProValSerProProIle 219
Db 760 TTGGAGGAGGAGCGCGCAAGCGCGCAAGCGCGCAAGCGCTGTGTGCGCACCCACATA 819
Qy 220 GluSerLysHisArgSerLeuValGlnIleLeuTyrAspGluLeuLeuArgLysLysLys 239
Db 820 GAATCAAGACATGAAAGCGCTGGTCCAGATCATCATGATGAGAACCGGAGAGAACCCGAA 879
Qy 240 AlaAlaHisArgLysLeuGluGlyProGlnValGluLeuProLeuTyrAsnGln 259
Db 880 GCGGCACACCGGCTCTCTAGAACGCTGGGGGCCCGCAGGTGGAGCTGCTCTGTACAACACAG 939
Qy 260 ProSerAspThrArgGlnTyrHisGluLeuLeuLysLysLysLysLysLysLysLys 279
Db 940 CCGTCTGACACACCGCAGTACCATGAAACATCAAAATATACCAAGCGGATGCGGAAGAG 999
Qy 280 LeuLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 299
Db 1000 CTGATCTGTGATCTTAAAGCGAGAGAACACCGCGCGCAGGAGTGGGAGACGCGCTTCTGC 1059
Qy 300 GlnArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgLysLysLysLys 319
Db 1060 CAGCGCTATGACCTAGCTCATGAGGCGGTGGGAGAGAGGTAGAGGCGCATAGAGAACCAAT 1119
Qy 320 ProArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGlu 339
Db 1120 CCGGAGAGGAGGCGCAAGGAGAGACAGGTGAGGAGTACTACCGAGAAACAGTTCCTCCGAG 1179
Qy 340 IleArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyClnArgGlySer 359
Db 1180 ATCCGAGAGCAGTGGGAGCTGACAGGAGCGCATGACAGAGCGGTGGCGCGCGTGGCAGT 1239
Qy 360 GlyLeuSerMetSerAlaAlaArgSerGluHisGluValSerGluLeuIleAspGlyLeu 379
Db 1240 GGGCTCTCCATGCGGCTGCCCGCCAGTGAAGCATGAGGTTCCTGAGATCATTCATGCGCTG 1299

Db 3397 ACACCCCGGCTACCGGCTCTGGGCTCTCCACGATAGTCCCGGCGCCGCTCTGCCA 3456
Qy 1098 ArgProThrIleSerAnProProProLeuIleSerSerAlaIleHisProSerVal 1117
Db 3457 CGTCCCGCC--ATCTCTAAACCCCGCCACCCCTCATCTCTCTGCAAGCATCCCGGGTA 3513
Qy 1118 LeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyr 1137
Db 3514 CTGTAGAGCGAGCTGGGTGCATCTCCAGAGGATGTCACTCAAGTTCGTGTGCTCAC 3573
Qy 1138 SerGluHisAlaIleAlaProValGlyProValThrMetGlyLeuProLeuProMetAsp 1157
Db 3574 TCAGAGCATGCCAAG--CCCATGGGCTCTCACCATGTGAGCTGCCCTTGGCGGTGAC 3630
Qy 1158 ProLysLysLeuAlaProPheSerGlyValLysGlnGlnLeuSerProArgGlyGln 1177
Db 3631 CCTAAGAGCTG----- 3642
Qy 1178 AlaGlyProGluSerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGly 1197
Db 3643 -----GGG 3645
Qy 1198 ThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSerThrArgVal 1217
Db 3646 ACAGCACT--GGCTCCGCCACACAGTGGAGCATCACCAAGGGCT--CCCATGATCCCGGGCT 3703
Qy 1218 ProSerAspSerAlaIleThrTyArgGlySerIleThrHisGlyThrProAlaAspVal 1237
Db 3704 GCAGAGCGGCC--AGCTACAGAGGCTCTATCACCCAGC--AGCGCCGAGAGCTC 3756
Qy 1238 LeuTyLysGlyThrIleThrArgIleIleGlyLeuAspSerProSerArgLeuAspArg 1257
Db 3757 CTCTAAGGTGTACATCAGAGATCGTGTGAGAGCAGCCCAAGTGTGCTTGACCGG 3816
Qy 1258 GlyArgGluAspSerLeuProLysGlyHisValIleTyGluGlyLysGlyHisVal 1277
Db 3817 GCACGAGAGGACACCTCGCCCAAGGGCCATGTCTATCATGAGGGCAAGAAAGGCGACGTC 3876
Qy 1278 LeuSerTyGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSer 1297
Db 3877 CTATCTATGAAGTGTATGTCGTGTACAGTGTCTAAGGNGGATGGAAGGAGCAGC 3936
Qy 1298 SerGlyProProHisGluThrAlaIleProLysArgThrTyArgMetMetGluGlyArg 1317
Db 3937 TGGGGCCACCCCATGAGACTGGCCCTTAACGACCTATGACATGATGAGGGCCGT 3996
Qy 1318 ValGlyArgAlaIleSerSerAlaSerIleGluGlyLeuMetGlyArgAlaIleProPro 1337
Db 3997 GTAGGACGAGCTGTCACTCAGCCAGCATAGAGGAGCTCATGGGCGCGGCCATC--CCT 4053
Qy 1338 GluArgHisSerProHisLysLysGlnHisHisIleArgGlySerIleThrGln 1357
Db 4054 GAGCAGCAGACCC--CAGCTCAGAGGAGCAGCATCACATCCGAGGCTCCATCAGCAG 4110
Qy 1358 GlyIleProArgSerTyValGluAlaGlnGluAspTyLeuArgArgGluAlaLysLeu 1377
Db 4111 GGCATCCCGAGTCTATGTGAGGCGCAGAGAGCTACTTACGGGGGAGGCCAAGCTC 4170
Qy 1378 LeuLysArgGluGlyThrProProProProProSerArgAspLeuThrGluAlaTy 1397
Db 4171 TTGAAGCAGAGAGGAGCAGCACCCTCCCGCCACCCAGCCTCGGAGCTGAGTACGACCTAC 4230
Qy 1398 LysThrGln-----AlaLeuGlyProLeuLysValLeuProAlaHisGluGlyLeu 1414
Db 4231 AAGCCCGGCGCCCTGGACCTCTGGGTCCCTGAAGTGTGAGCCGAGCTCAGAGGGGTG 4290
Qy 1415 ValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArg 1434
Db 4291 GTAGCAACTGTGAAGAGGCGGCGCTCTATCATGATCCGAGAGAGAGCTGCGC 4350
Qy 1435 HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly 1454
Db 4351 CGCACACCTGAGCTACCCCTGGCAGCAGCGCTCTGAAGAGGGGTTCATCACCCAGGGC 4410

Qy 1455 ThrProLeuLysTyArgThrGlyAlaSerThrThrGlySerLysLysHisAspValArg 1474
Db 4411 ACCCCACTCAATACGACTCTGGGGCAGCCCTCCACTGGCAGCAAGAAACACGACGTGGC 4470
Qy 1475 SerLeuIleGlySerProGlyArgThrPheProProValHisProLeuAspValMetAla 1494
Db 4471 TCCATCATCGGAGGCGCGCGGCTTTCCTGCTGCTGCGCCCTGGACATATAGGCT 4530
Qy 1495 AspAlaArgAlaLeuGluArgAlaCysTyGluGluSerLeuLysSerArgProGlyThr 1514
Db 4531 GAGCGCGGGCTGTGAGCGGTGCTGTATGAGAGAGTCTGAAGAGCGCGGTGAGGAC 4590
Qy 1515 AlaSerSerSerGlySerIleAlaArgGlyAlaProValIleValProGluLeuGly 1534
Db 4591 AGCAGTGTGCTGGGGCTCCATCACAGTGGGGCTCCAGTGTGCTGCTGAACCTGGC 4650
Qy 1535 LysProArgGlySerProLeuThrTyGluAspHisGlyAlaProPheAlaGlyHisLeu 1554
Db 4651 AAGCCAGGCAAGCCCACTGACTTACGAGAGCAGCGGCGCCCTTCACAGTACCTG 4710
Qy 1555 ProArgGlySerProValThrMetArgGluProThrProArgLeuGlnGluGlySerLeu 1574
Db 4711 CCACGTGGCTCCCTGTGACGAGGAGGAGCCAGCCAGCCCTTCAGGAAGCAGCCTC 4770
Qy 1575 SerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProArgGluIleAlaLys 1594
Db 4771 CTATCCAGCAGAGGCTCCAGGACCGGAAGCTGACATCTACACCCCGGAGAGTCCGCAAG 4830
Qy 1595 SerProHisSerThrValProGluHisHisProHisProHisProTyGluHisLeu 1614
Db 4831 TCCCCACAGAGCTGTGCGCGGAGCACCCTCACCCATCTCCCTATGAGCACTTG 4890
Qy 1615 LeuArgGlyValSerGlyValAspLeuTyArgSerHisIleProLeuAlaPheAspPro 1634
Db 4891 CTCGGGGGCTGTGTGTGTGACCTGTATCCGTGTGCATATCCCTATGGCTTTGACCC 4950
Qy 1635 ThrSerIleProArgGlyIleProLeuAsp---AlaAlaAlaAlaTyTyLeuProArg 1653
Db 4951 ACCTCCATACCCGAGGGATCCCTCTGGAAGCAGCAGCGCAGCTACTACCTGCCCCG 5010
Qy 1654 HisLeuAlaProAsnProThrTyProHisLeuTyProProTyLeuIleArgGlyTy 1673
Db 5011 CATTTGGCCCCAGAGCCCTTACCCACCTGTACCCACCTTACCTCATCTCCGCGGTAC 5070
Qy 1674 ProAspThrAlaAlaLeuGluAsnArgGlnThrIleLeuAsnAspTyIleThrSerGln 1693
Db 5071 CTGTACACCGCGGCTGTGAGAACCCGACAGCATCATCAATGACTACATCATCTCGCAG 5130
Qy 1694 GlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeu 1713
Db 5131 CAGATCACCAACAGCTGCTCCGCGATGGCCAGCGTGTGCATGCTGAGGGGTCTG 5190
Qy 1714 SerProArgGlySerSerLeuAlaLeuAsnTyAlaAlaGlyProArgGlyIleLeuAsp 1733
Db 5191 TCACCGGAGAGTCTCGCTGGGCTTCAATATTCCGCTGGCCCAAGAGGCAATATGAC 5250
Qy 1734 LeuSerGlnValProHisLeuProValLeuValProThrProGlyThrProAlaThr 1753
Db 5251 CTGTCCCAAGTCCACACCTGCGGTGTGTGTGTCACCAAGCAGCGGACCCCTGCGCAC 5310
Qy 1754 AlaMetAspArgLeuAlaTyLeuProThrAlaProGlnProPheSerSerArgHisSer 1773
Db 5311 GCATCGACCGCTTGTGCTTACTCCCTGCGGCCCCCACTTCAAGCAGCGCCACAGT 5370
Qy 1774 SerSerProLeuSerProGlyTyProThrHisLeuThrLysProThrThrThrSerSer 1793
Db 5371 AGTTCACCGCTTCCCGAGAGGCGCCACTCACTAGCTAAACCACTGCCCATCTTCA 5430
Qy 1794 SerGluArgGluArgAspArgGluArgAspArgAspArgGluArgGluLysSer 1813
Db 5431 TCGAGCGGAGACGGGAACGTGAGCGGGAACGAGAC-----AAGTCC 5472

QY 1814 IleLeuThrSerThrThrThrValGluHisAlaProIleThrArgProGlyThrGluGln 1833
DB 5473 ATCTCTACACGTCTACACGTACAGTGGAGCATGACCATCTGGAGACCTGGTACGGAGCAG 5532
QY 1834 SerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerArgProAlaSer 1853
DB 5533 AGCAGCGGGGCT-----GGGGGCAGCAGCGCGCGCGCTCC 5568
QY 1854 HisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnHisAlaLeuGln 1873
DB 5569 CACACC-----CACAGCAGCTGGCCCATCTCCCGGACCCAGACGCTTGCAGCAG 5622
QY 1874 ArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaValGluProSer 1893
DB 5623 AGGCCCGAGTGTCTGCACACACAGCAGCATGAAGGGCGTGTCACTCCGTGGAAACCCGCG 5682
QY 1894 LysProThrValLeuArgSerThrThrSerSerProValArgProAlaAlaThrPhe 1913
DB 5683 ACGCCCAACCGTCTGAGGTCCACCTCCACCTCTCGCTGCTCGCCAGCTGCACATTC 5742
QY 1914 ProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyProThrLeuMet 1933
DB 5743 CCACCTGCCACCCACTTGGTGGCACCTTGAAGGGGTCTACCTTACCTTACCTCATG 5802
QY 1934 GluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProArgAla 1953
DB 5803 GAGCCCGTCTCTGTACCCAGAGAGCTCTCGGGTGGCGCCCGCCGAGCGGCCCGGTGTG 5862
QY 1954 AspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAlaSer 1973
DB 5863 GAGGTGGCCATGCTCTCTCCACCAACCCCGCGCGG-----GAGCCCGCTCC 5913
QY 1974 SerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAlaThr 1993
DB 5914 TCACCCAGCAAGAGCTCCGAGCCCGCATCTCCAGACCCCGAGCTCCAGCCACACAGCC 5973
QY 1994 IleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProAla 2013
DB 5974 ATGCCCGCACCCAGCAAGAGCTTGCACCCACCCACCATGCTCGAGTCGGACCGCGCGGG 6033
QY 2014 ProProAlaSerAlaSerAspProHisArgGlyThrGlnSerLysProPheSerIle 2033
DB 6034 ---CCACCTCGGCTCAGATCTGCACCCGAGAAAGACTCAAGTAAACCTTTTCCATC 6090
QY 2034 GlnGluLeuGluLeuArgSerLeuGlyTyHis---GlySerSerTySerProGluGly 2052
DB 6091 CAGGAATTGGAACTCCGTTCTCTGGGTACCAACAGTGGAGTGGCTACAGCCCGCATGGG 6150
QY 2053 ValGluProValSerProValSerProSerLeuThrHisLysAlaLysGlyLeuProLys 2072
DB 6151 GTGGAGCCCATCAGCCCGGTGAGCTTCCCGAGCTTCCCGAGCTGACCAAGAGGGGTCTCCAAA 6210
QY 2073 HisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProGly 2092
DB 6211 CCTCTGGAAGAGCTAGAGAAGAGCCACTTGGAAAGGGAGCTGGCGCACAGCAGCGAGCC 6270
QY 2093 ProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLysLeuArgProLeuProGluSer 2112
DB 6271 CCCATGAAGCTCAGCGCGGAGGTGCGCATCTCCACATCTCGGGCCACTGCGCCGAGAGC 6330
QY 2113 GlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgVal 2132
DB 6331 CAGCCCTCATCCAGCCCATCTCTCCAGATGTGCCCGAGCTCAAAAGTCAACAGAGGGTG 6390
QY 2133 ValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyThrArgHisPro 2152
DB 6391 GTCACCTCTGGCTCAGCATCATCAGCAGAGGTCAATAGCAGAGGTATACACCCGCGACCCCG 6450
QY 2153 GlnGlnLeuSerAlaProLeuProAlaProLeuTySerPheProGlyAlaSerCysPro 2172
DB 6451 CAGCAGCTCAGTGGGCCCTTCCCGCCCTCTCTACTCTCTTTCCGGAGCCAGCTGCCCT 6510
QY 2173 ValLeuAspLeuArgArgProProSerAspLeuTyLeuProProAspHisGlyAla 2192

DB 6511 GTGTGTGATCTTCGGCGGCCACCCAGTGACCTTACTCCACCCCGCCACCATGGCACC 6570
QY 2193 ProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAsnLysThr 2212
DB 6571 CCAGCCCGGGGTTCCTCCCCACAGTGAAGGGGGCAAAAGGTCTCCAGAACCCAGCAAAACA 6630
QY 2213 SerValLeuGlyGlyGlyGluAspGlyIleLeuProValSerProProGluGlyMetThr 2232
DB 6631 TCGGTCTCTGGGAGAGCAGTGTGAGATGCCATGTGACCTGTGTATCCACAGAGGGGTACT 6690
QY 2233 GluProGlyHisSerArgSerAlaValTyProLeuLeuTyArgAspGlyGluGlnThr 2252
DB 6691 GAGCCAGGACATGCTCGGAGGGCTGTGTACCCATGCTGTATCCAGAGCGGGAAACAGGC 6750
QY 2253 GluProSerArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePhe 2272
DB 6751 GAGCCC---AGATGGGCTCTAAGTCTCCAGGCAACACCCAGCCAGCGCCCTTCTTC 6807
QY 2273 SerLysLeuThrGluSerAsnSerAlaMetValLysSerLysGlnGlnIleAsnLys 2292
DB 6808 AGTAAGCTGACGAGAGCACTCCGCCATGTGTGAAGTCGAAGAGCAGGAGATCAACAG 6867
QY 2293 LysLeuAsnThrHisAsnArgAsnGluProGluTyAsnIleSerGlnProGlyThrGlu 2312
DB 6868 AAATCTCAACACACACACCGGAAACGAGCCAGAAATACAATATTGGCCAGCTGGGACGAA 6927
QY 2313 IlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyArgSerGlnAlaVal 2332
DB 6928 ATCTTCAACATCCCGCCATCATCTGAGCAGGCTTATGACCTGTAGAACCCAGCGGTG 6987
QY 2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352
DB 6988 CAAGAACACGCGAGCAGCAACCAATGGGCTAGAGGCCATTTATTAGAAAGGCACATCGGGT 7047
QY 2353 LysTyAspGlyTrpGluLeuSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372
DB 7048 AAATATGATCATTTGGAAAGAGCCCGCGCTCGGGCCCAATGTCTTTAAACCTCTCAAT 7107
QY 2373 AlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyArgSerAsp 2391
DB 7108 GCCAGCCCGAGCTGCGCGCTGTCTGTATGCCCATTAACCATGTGTAGCAGCGAGTGAC 7167
QY 2392 HisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArgProSerSer 2411
DB 7168 CACGCACTCATCTCCCGAGGTGGAGTGGGAAGCCAAAGTCTCTGGCAGACCTAGCAGC 7227
QY 2412 ArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerValSer 2431
DB 7228 CGAAGGCCCATTCGCCAGCACCAGGCTAGCGTCCGGAGCCGACCCCTTCTGTCTCC 7287
QY 2432 SerValHisSerGlyLysAspCysAsnArgArgThrProLeuThrAsnArgValTrpGlu 2451
DB 7288 TCAGTACATCTCAGGGGAGTGCATATCGCCGAACACCACTCACCAACCGGTGTGGAG 7347
QY 2452 AspArgProSerSerAlaGlySerThrProPheProTyAsnProLeuIleMetArgLeu 2471
DB 7348 GACCGGCCCTCTCTGCGAGGTCACGCAATCCCTCAACCCCTTTGATTATGAGGCTA 7407
QY 2472 GlnAlaGlyValMetAlaSerProProProGlyLeuProAlaGlySerGlyProLeu 2491
DB 7408 CAGCGAGGTGTATGGCTCTCCCGCCACCTGGCTTGGCGCAGCAGCGGGCCCTTA 7467
QY 2492 AlaGlyProHisHisAlaTrpAspGluProLysProLeuLeuCysSerGlnTyGlu 2511
DB 7468 GCTGGTCCCCCAGCGCTGGATGAGAGCCCAAGGACCTGTGTGTTCACAGTATGAG 7527
QY 2512 ThrLeuSerAspSerGlu 2517
DB 7528 ACATCTCCGAGAGCGAG 7545

QY 281 IleLeuTyrPheLysArgGlnHleAlaArgLysGlnTrrLysGlnLysPheCysGln 300
DB 546 ATCTTGTACTTTAAGCGGAGGACCAACCGCGCGCAAGCAAGTGGAGACAGCGCTTTCGCG 605
QY 301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgLysGlnLeuAsnPro 320
DB 606 CGTATATGACCGCTCATGAGGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 665
QY 321 ArgArgArgAlaLysGlnLysValArgGlnTrrLysGlnLysGlnPheProGluLeu 340
DB 666 CGAAGGAGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 725
QY 341 ArgLysGlnArgGlnLeuGlnLysGlnMetGlnSerArgValGlnArgGlnLysGln 360
DB 726 CGCAAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 785
QY 361 LeuSerMetSerAlaAlaArgSerGlnHleGlnValSerGlnLysLysLysLysLysLys 380
DB 786 CTCTCCATCTCGGCTGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 845
QY 381 GluGlnGlnLeuGlnLysGlnMetArgGlnLeuAlaValLysLysLysLysLysLysLys 400
DB 846 GAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 905
QY 401 AspAlaAspGlnArgLysPheLysPheLysPheLysPheLysPheLysPheLysPheLys 420
DB 906 GAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 965
QY 421 LysValTyrLysAspArgGlnValMetAsnMetTrrSerGlnGlnGlnGlnGlnGlnGln 440
DB 966 AAGGCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1025
QY 441 ArgGlnLysPheMetGlnHleProLysAsnPheGlyLysLysLysLysLysLysLysLys 460
DB 1026 CGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1085
QY 461 LysThrValAlaGlnCysValLeuTyrTyrLeuThrLysLysLysLysLysLysLysLys 480
DB 1086 AAGCGGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1145
QY 481 SerLeuValArgArgSerTyrArgArgArgLysSerGlnGlnGlnGlnGlnGlnGlnGln 500
DB 1146 AGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1205
QY 501 Gln 520
DB 1206 CAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1265
QY 521 AspGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGln 540
DB 1264 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1313
QY 541 AspLysGlnAspLeuLeuLysGlnLysLysLysLysLysLysLysLysLysLysLysLys 560
DB 1314 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1373
QY 561 LysGlnAlaValAlaSerLysGlnLysLysLysLysLysLysLysLysLysLysLysLys 580
DB 1374 AAGAGGCGGCTGCTCCAAAGGCGCGCAAACTGCAACAGCAGGCGCGCGCGCGCGCG 1433
QY 581 ArgLysThrArgSerMetAlaAsnGlnAlaAsnSerGlnLysLysLysLysLysLysLys 600
DB 1434 CGTATCACCGGCTTCATGCGCAACGAGGCGCAACCATGAGGAGGAGGAGGAGGAGGAG 1493
QY 601 SerAlaGlnLeuAlaSerMetGlnLeuAsnGlnLysSerArgTrrThrGlnGlnGlnGln 620
DB 1494 AGTTTCAGGCTGCTTCCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1553
QY 621 GluThrAlaLysLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 640
DB 1554 GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1613
QY 641 ValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArgGln 660

DB 1614 GTGGGCTCCAAAGCAACCGTGTCCAGTGTAAAGACTTCTACTTCAACTACAGAGAGGAG 1673
QY 661 AsnLeuAspGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 680
DB 1674 AACCTGAGCAGATCTCTTCAGCAGCAGCAAGCTAAGATGAGAGAGAGAGAGAGAGAG 1733
QY 681 ArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 700
DB 1734 AGGAG 1793
QY 701 GluAspGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 720
DB 1794 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1853
QY 721 AlaGlnAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 739
DB 1854 GCAGAGAGGCTCTAGGAGGCTCTGGGATGAGGTTCCAGAGTTGGGAGTTGGGAGTTGG 1913
QY 740 AlaThrValAsnAsnSerSerAspThrGlnSerLysLysLysLysLysLysLysLysLys 759
DB 1914 GCTGCTGTCAAGCAACAGCTCTCATCTAGAGTGTCCATCCCGCGTTCAGAGAGCAGC 1973
QY 760 LysAspThrGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 777
DB 1974 AAGGAGCACT-----GGGCTTAAACCACTGGGCACTGAAGCACTGCCGCTGCCACC 2024
QY 778 ProProProGlnProProProThrProProArgThrSerArgAlaProLysGlnProThr 797
DB 2025 CAGCCAGCTGTGCT-----CCTCCAGAAAGAACCGGAGTAGGAGGAGGAGGAGGAGG 2078
QY 798 ProAlaSerGlnAlaThrGlyAlaProThrProProProAlaProProProProSerAla 817
DB 2079 CCAGTCCCTGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2134
QY 818 ProProProValValProLysGlnLysGlnLysGlnLysLysLysLysLysLysLysLys 837
DB 2135 CCCCAGGCTACTGTGAGCAAGGATGAACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2194
QY 838 GluGlnLysGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 857
DB 2195 GAGATGCCAAG 2245
QY 858 Ala-GluGlnLys-----ValLysSerGlnCysThrGln 868
DB 2246 GCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2305
QY 868 uGluAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 887
DB 2306 GGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2365
QY 887 aGluGlnAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 907
DB 2366 TGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2422
QY 907 sSerSerGlnValProGlnAspSerAspSerSerSerSerSerSerSerSerSerSerSer 927
DB 2423 CAGCTCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2482
QY 927 pGluAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 947
DB 2483 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2542
QY 947 oThrGlyAspProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 967
DB 2543 GCGTGGAGATCTCCCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2602
QY 967 sGlnArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 987
DB 2603 GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2659
QY 987 uAspAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 1007

Db	2660	GGACACAGTACCCCCAAAGCCAGTTGTTCCCTCTGGCTCCACCCACGACGACCTACAGCC	2711
Qy	1007	odluserAspAlaProGlnInProGlySerProArgGlyLysSerArgSerProAl	1027
Db	2720	AGAGGGTGAAGTGTCTCAGCAGTGGGAGGAGTCCAGTGGCAAGTCCGCGACGCCAGT	2779
Qy	1027	apProAlaAspLysGlu-----AlaPhealaAalaGluAl	1039
Db	2780	GCCTCTGCCGAGAAAGGCGAGAAACCGCATCTCTCCGGCTTCCACCTGAGGG	2839
Qy	1039	aGlnLysLeuProGlyYaspProProCyfThrThrSerGlyLeuProPheProValProPr	1059
Db	2840	CCCAAGACTACCACTGAGAGCCCAACGCTGGTCATCGGGCCCTGCCCTTCCCATCCCTCC	2899
Qy	1059	oArgGluValLysAlaSerProHisAlaProAspProSerAlaPheSerTyrAlaPr	1079
Db	2900	ACGGAGAGTGATCAGACTTCCCAACGCGCGTCAACCCCTCTGCTCTCTCTACACACC	2959
Qy	1079	oProGlyHisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArgPr	1099
Db	2960	CCCCGGTCAACCGCTGCTTGGGCTTCCAGTAGTGGCGGCGCGTCTTGCACAGTCC	3019
Qy	1099	oProThrLysSerAsnProProProLeuLysSerSerAlaHisProSerValLeuGl	1119
Db	3020	CCCC---ATCTTAACCCCCCAACCCCTCATCTCTCTGCCAAGCATCCCGGGGTACTTGA	3076
Qy	1119	uArgGlnLysGlyAlaLysSer---GlnGlyMetSerValGlnLeuHisValProTyrSe	1138
Db	3077	GAGGCAGCTGGGTGCCATCTCCAGCAGGAGGATGTCACTCCAGCTGCTGTGCCTCACTC	3136
Qy	1138	rGluHisAlaLysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspPr	1158
Db	3137	AGAGCTGCCAGGCCCCCTTGGGCTCTTCACTATGGGGCTGCCCTTCCGCTGGAGACC	3196
Qy	1158	oLysLysLeuAlaProPheSerGlyValLysGlnGlnLeuSerProArgGlyGlnAl	1178
Db	3197	TARGAAGCTG-----3206	3206
Qy	1178	agLysProGlySerLeuGlyValProThrAlaGlnGluAlaSerValLeuArgGlyTh	1198
Db	3207	-----GGGAC3211	3211
Qy	1198	rAlaLeuGlySerValProGlyGlySerLysThrLysGlyLeProSerThrArgValPr	1218
Db	3212	AGCACTGGGCTCGCCACCATGAGAGCATCAACAGGGGCTTCCCCAGTACCCGGGTGC	3271
Qy	1218	oSerAspSerAlaLysThrTyrArgGlySerLysThrHisGlyThrProAlaAspValLe	1238
Db	3272	AGACGGCCCC---AGCTACAGAGGCTCTATCAACCCACGACCGCCCGCAGACGTCCT	3325
Qy	1238	uTyrLysGlyThrLysThrArgLysLysGlyLysValLysGlyHisValLe	1258
Db	3326	CTACAGGGTACCNCAGCAGGATCGTGGTGAGCAGACCCCAAGTCTGTACCGGGC	3385
Qy	1258	YArgGluAspSerLeuProLysGlyHisValLysGlyLysGlyHisValLe	1278
Db	3386	ACGAGAGACACCTTGCACAGGAGGATGATCATCTATGAGGGCAGAAAGGCCAGTCCT	3445
Qy	1278	uSerTyrGluGlyGlyMetSerValThrGlnCysSerLysGluAspGlyArgSerSerSe	1298
Db	3446	ATCTCATGAAGTGGTATGTCGCTGTCACAGTGTCTTAAGGAGGATGGAAGGAGCAGCTC	3505
Qy	1298	rGlyProProHisGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgVa	1318
Db	3506	GGGCCCCACCCATGAGACTGCGCGCCCTTAACGCACCTATGACATGATGAGGGCCGTGT	3565
Qy	1318	lGlyArgAlaLysSerSerAlaSerLysGlyLeuMetGlyArgAlaLysProProGl	1338
Db	3566	AGGCAGAGTGTCACTCAGCAGCATAGAGGGATCTCATGGGCGCGGCCATC---CCTGA	3622
Qy	1338	uArgHisSerProHisHisLeuLysGluGlnHisLysLysLysLysLysLysLysLysLys	1358
Db	3623	GCAGCAGCCCC---CACTCTAAGAGAGAGATCATCATCCAGGCTTCAACAGCGCAAGG	3679

Qy	1358	YlleProAigS	rTyValGluAlaGlnGluAspTyxLeuArGArGluAlaLysLeuLe	1378
Db	3680	CATCCCGAGGT	CTATGTGGAGGCGCAGAGGAGCTACTTACGGCGGGAGGGCCAGACTCTT	3739
Qy	1378	uLysArGgGluG	ThrProProProProProProSerArGAspLeuThrGluAlaTyLy	1398
Db	3740	GNAGCGAGAGG	AGACACACACACACCCACCACCACTCGGGACCTGACTGAGACTCAAA	3799
Qy	1398	sThrGln----	-----AlaleuGlyProLeuLysLeuLysProAlaHisGluGlyLeuVa	1415
Db	3800	CCCCCGCCCC	TGGACCTCTGGGTCCCTTGAGCTGAGCCGACTCAGAGGGGTGGT	3859
Qy	1415	AlaThrValHis	GluAlaGlyArgSerIleHisGluIleProArGgGluGluLeuArGHi	1435
Db	3860	AGCACTGTGA	GGAGCGGGCGGCTCTATCCATGAGATCCCGAGAGAGAGCTGCGCCG	3919
Qy	1435	sThrProGluLeu	ProLeuAlaProArGProLeuLysGluGlySerIleThrGlnGlyTh	1455
Db	3920	CACACTGTAGC	TACCCCTGGCACACCGGCTCTGAAGAGGGTTCCATCACCCAGGGCAC	3979
Qy	1455	rProLeuLysTy	AspThrGlyAlaSerThrGlySerLysLysHisAspValArgSe	1475
Db	3980	CCNCTCAAGT	AGNCTCTGGGCGACCTTCCACTTGGCACCCAGNAACACGAGTGCCTC	4039
Qy	1475	rLeuIleGlySer	ProGlyArgThrPheProProValHisProLeuAspValMetAlaAs	1495
Db	4040	CATCATCGGCA	CCCCCGGCGGCTTCTCCCTGCGCCTGCACCGCTGGACATAATGGCTGA	4099
Qy	1495	pAlaArgAlaLeu	GluArgAlaCysTyxGluGlySerLeuLysSerArgProGlyThrAl	1515
Db	4100	CGCCCGGGCAC	TGAGAGGTGCCTGTGTATGAAGAGAGTCTGAAGAGCGGCTCAGGACCA	4159
Qy	1515	aSerSerSerG	lyGlySerIleAlaArgGlyAlaProValIleValProGluLeuGlyLy	1535
Db	4160	CAGTGTGCGA	GGGCTCCATCACAGTGGGGCTCCAGTGTGTGCTGACTGAGCTGGGCA	4219
Qy	1535	sProArgGlnSer	ProLeuThrTyxGluAspHisGlyAlaProPheAlaGlyHisLeuPr	1555
Db	4220	GCACCGGCAAA	CCCATCTGACTTACGAAGACACGGGGCACCTTCCACAGTCACTGGCC	4279
Qy	1555	oArgGlySerP	roValThrMetArgGluProThrProArGLeuGlnGluGlySerLeuSe	1575
Db	4280	ACGTGTGCTCC	CTGTGTACCAACAGGAGAGCCCAACGCCAGCCTTCAGGAAGGACGCTCT	4339
Qy	1575	rSerSerLysAla	SerGlnAspArgLysLeuThrSerThrProArGgGluIleAlaLysSe	1595
Db	4340	ATCCACAGGGT	CCCCGAGCCGGAGCTGACATCTACACCCCGGGAGATCGCCAGTCT	4399
Qy	1595	rProHisSerThr	ValProGluHisHisProHisProIleSerProTyxGluHisLeuLe	1615
Db	4400	CCACACAGCAT	GTGGCCCGAGACACCCCTCACCCCATCTCCCTATGAGCACTGGCT	4459
Qy	1615	uArgGlyValSer	GlyValAspLeuTyxArgSerHisIleProLeuAlaPheAspProTh	1635
Db	4460	CCGGGCGGTGA	TGGTGTGAGCTGTACCGTGGTGCACATCCCATTTGGGCTTTGACCCCA	4519
Qy	1635	rSerIleProArg	GlyIleProLeuAsp---AlaAlaAlaAlaTyTyxLeuProArGHi	1654
Db	4520	CTCCATACCCG	AGGGATCTCTTGGAGAGCAGCGGACGCTACTACTCTGCCCGGCA	4579
Qy	1654	sLeuAlaProLeu	ProThrTyxProHisLeuTyxProProTyxLeuIleArgGlyTyxPr	1674
Db	4580	CTTGGCCCCCA	CCCCACCTACCCACACTGTACCCCATCTTACTCTATCGCGGCTACCC	4639
Qy	1674	oAspThrAlaAla	LeuGluAspArgGluThrIleIleAspAspTyxIleThrSerGlnI	1694
Db	4640	TGACAGCGGGT	CTTGAGAACCGGACACCATCATCATGACTACATCACTACCTCGGAGCA	4699
Qy	1694	nMetHisHisAsp	ThrAlaThAlaMetAlaGlnArgAlaAspMetLeuArgGlyLeuSe	1714
Db	4700	GATGCACCACT	CTGCTCTCGGCATCGGCCAGCGTGTGACATGCTGAGGGGCTGTGCT	4759

QY 1714 rProArgGluSerSerLeuAlaLeuAenTyrAlaAlaGlyProArgGlyIleAlaSerLe 1734
DB 4760 ACCGCGAGAGTCTCGCTGGCCCTCAATTATGCGCTGGCCCAAGAGGCATTTATCGACCT 4819
QY 1734 uSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaThrAl 1754
DB 4820 GTCCCAAGTGGCCACACCTGCGCGGTGTGTGTGCACCAACGCCAGGCACCCCTGCCACCGC 4879
QY 1754 aMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisSerSe 1774
DB 4880 CATCGACCGCTTGCCTACCTCCGCCCTGCGGCCACCCCTTCAGCAGCGCCACAGTAG 4939
QY 1774 rSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrThrThrSerSerSe 1794
DB 4940 CTCACCGCTGTCGCCAGGAGGCCCACTCACCTAGCTAAACCACTGCCACATCTTCATC 4999
QY 1794 rGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluArgLysSerIl 1814
DB 5000 GGAGCGGAGAACGGACCTGAGCGGAGACGAGAC-----AAGTCCAT 5041
QY 1814 eLeuThrSerThrThrValGluHisAlaProIleTrpArgProGlyThrGluGlnSe 1834
DB 5042 CCTCAGCTCTACCACTACAGTGGAGCATGACCCCATCTGGAGACCTGGTACGGAGCAGAG 5101
QY 1834 rSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerArgProAlaSerHi 1854
DB 5102 CAGCGGGGCT-----GGGGGCGAGCAGCGCGCGCCCTCCCA 5137
QY 1854 sSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAlaLeuGlnLar 1874
DB 5138 CACC-----CACCAGCAGCTCGCCCATCTCCCGCCGACCCAGAGCGCTTGCAGCAGAG 5191
QY 1874 gProSerValLeuHisAnThrGlyMetLysGlyIleIleThrAlaValGluProSerLy 1894
DB 5192 GCCCAGTGTGTGCACACACGAGCATGAAGGCGTGTCTACCTCGCTGGAACCCCGGCAC 5251
QY 1894 sProThrValLeu-----ArgSerThrSerThrSerSerProValArgProAlaAl 1911
DB 5252 GCCCAGCGTCTGAGTGGGCGCAGTCCACTTCCACTTTCGCTGTGCGGCCAGCTGC 5311
QY 1911 aThrPheProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProTh 1931
DB 5312 CACATTCCTCACCCTGACACCTGCGCCACTTGGTGGCACCCTTGAAGGGGTCTACCTTAC 5371
QY 1931 rLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgPr 1951
DB 5372 CCTCATGAGCGCGTCTGTGTACCAAGGAGACCTCTCGGGTGGCGCGCGCGCGCGCGGC 5431
QY 1951 oArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluPr 1971
DB 5432 CCGGGTGGAGCGTGGCCATGCTTTCTTACCAACCCCGGGCGCG-----GAGCC 5482
QY 1971 oAlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHi 1991
DB 5483 CGCCTCTCACCACAGCAGAGCTCGAGCCCGCATCTTACCAACCCCGCGCGCGCGCGCGCG 5542
QY 1991 sAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspPr 2011
DB 5543 CACAGCCATCGCGCGCAGCCAGCAAGAACCTTCCACCCCACTGCGCGCGCGCGCGCGCG 5602
QY 2011 oProAlaProProAlaSerAlaSerProHisArgGluLysThrGlnSerLysProPh 2031
DB 5603 GCGGGG-----CCACCTCGCGCTCAGATCTGCACCGAGAAAGACTCAAGTAACCTT 5659
QY 2031 eSerIleGlnGluLeuGluArgSerLeuGlyTyrHis---GlySerSerTyrSerPr 2050
DB 5660 TTCATCCAGGAATGGAACTCGCTTCTCTGGGTATACCAAGTGGAGCTGGGTATACGCGC 5719
QY 2050 oGluGlyValGluProValSerProValSerProSerLeuThrHisAspLysGlyLe 2070
DB 5720 CGATGGGGTGGAGCCCATCAGCGCGGTGAGCTTCCCGCGCTGACCCAGCAAGAGGGCT 5779
QY 2070 uProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyLeuArgProLysG 2090

DB 5780 CTCCAAACCTCTGGAAGAGCTAGAGAGAGCCACTTGAAGGGAGGTGGCGCACAGCA 5839
QY 2090 nProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuPr 2110
DB 5840 GCCAGGCCCATTAAGCTCAGCGGGAGGCTGCCCATCTCCACATCTCGGGCCACTGCG 5899
QY 2110 oGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisG 2130
DB 5900 CGAGAGCAGCGCTCATCGCCCACTCTCCAGACTGCCCGCAGCATCAAGAGTCAACA 5959
QY 2130 nATGValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHi 2150
DB 5960 GAGGGTGGTCACTCTGGCTCAGCACATCAGCAGAGGTCAATACGCGAGACTACACGCCCA 6019
QY 2150 sHisProGlnGluSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSe 2170
DB 6020 CCACCCGCGAGCTCTCAGTGGCCCTTCCCGCCCTCTCTACTCTCTTCCCGAGCCAG 6079
QY 2170 rCysProValLalaAspLeuArgArgProProSerAspLeuTyrLeuProProPhePh 2190
DB 6080 CTGCCCTGTCTTGGATCTTCGCCGCCCACTGACCTCTACCTCCCACTCCCGCCCA 6139
QY 2190 sGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerProGluProAs 2210
DB 6140 TGGCACCACGAGCTCGGGCATCCGCCCACTGAGGGGGGCAAAAGGTCCCGACAGCCAG 6199
QY 2210 nLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerProGluGl 2230
DB 6200 CAAACATCGCTCTGGGCACGAGGATGCCATTAGCTGTGTCCCAACAGAGGG 6259
QY 2230 yMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrArgAspGlyGl 2250
DB 6260 CATGACTGAGCGAGGACATGCTCGAGCAGCTGCTACCCACTGTGTATCGAGCGGGA 6319
QY 2250 uGlnThrGluProSerArgMetGlySerLysSerProGlyAenThrSerGlnProAl 2270
DB 6320 ACAGGCGAGCG-----AGGATGGGTCTAGAGTCTCCAGGCAACACAGCGCGCGCAAC 6376
QY 2270 aPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysGlnGluI 2290
DB 6377 CTCTTCAGTATCTGCTGACTGAGACAACTCCGCCATGTGTGAAGTCGAAGACGAGAT 6436
QY 2290 aEnLysLysLeuAlaThrHisAenArgAenGluProGluTyrAsnIleSerGlnProGl 2310
DB 6437 CAACAAGAACTCAACACCCCAACCGGAACGAGCCAGAAATACAATATTGGCCAGCTGG 6496
QY 2310 YThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerGl 2330
DB 6497 GACGGAATCTTCAACATGCCCGCATCTAGGAGCGGCTTATGACCTGTAGAAGCCA 6556
QY 2330 nAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleArgLysAlaLe 2350
DB 6557 GCGGGTGAAGAACAGCCAGCAGCAACATGGGGCTAGAGCCCATTTATGAAGGCAT 6616
QY 2350 uMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAenAlaPheAenPr 2370
DB 6617 CATGGGTAAATTTGATCAGTGGGAAGAGCCCGCGCTCGCGCGCAATGCTTTTAAACCC 6676
QY 2370 oLeuAenAlaSerAlaSerLeuPro---AlaAlaMetProIleThrAlaAlaAspGlyAr 2389
DB 6677 TCTGAATGCCAGCGCCAGCTGTGCTGTATGCCCATTAACCTCTCTGACCGAGC 6736
QY 2389 gSerAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArgPr 2409
DB 6737 GAGTGACCACTGCTCCTCGCCAGTGGAGTGGGAAGCCCAAGGTCTCTGGCAGACC 6796
QY 2409 oSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSe 2429
DB 6797 TAGCAGCCGAATAGGCAAGTCCGAGCAGCAGCGCTAGCGTCCGAGAGACGACCCCTTC 6856
QY 2429 rValSerSerValHisSerGluGlyAspCysAenArgArgThrProLeuThrAenArgVa 2449

Db 6857 TGCTCTCTCAGTACACTCAGAGGGGAGCTGCAGTCCGCGGACCAACCACTCACCACCCCTGT 6916

Qy 2449 lTrpGluAspArgProSerSerAlaGlySerThrProPhaProTyAsnProLeuIleMe 2469

Db 6917 GTGGGAGACCGGCGCTCATCTCAGGGTCCAGCCATTCCTCCCTCAACACCTTTGATTAT 6976

Qy 2469 tArgLeuGlnAlaGlyValMetAlaSerProProProGlyLeuProAlaGlySerG 2489

Db 6977 GAGGCTACAGGAGGTGTCTATGGCTCCCGCGCCCTGGCTTGGCGAGGAGCGG 7036

Qy 2489 yProLeuAlaGlyProHisAlaTTrpAspGluLupProLeuLeuCySerG 2509

Db 7037 GCCCTAGTCTGCTCCACCACTGGGATGAGAGCCAGCCACTGCTGTGTTCACA 7096

Qy 2509 nTyrgluThrLeuSerAspSerGlu 2517

Db 7097 GTATGAGACACTCTCGGACGAG 7121

RESULT 6

HSU37146

LOCUS

DEFINITION

5989 bp mRNA linear PRI 31-OCT-1995

Human silencing mediator of retinoid and thyroid hormone action

ACCESSION

U37146

VERSION

U37146.1 GI:1045654

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 5989)

Chen, J.D. and Evans, R.M.

A transcriptional co-repressor that interacts with nuclear hormone receptors

Nature 377 (6548), 454-457 (1995)

96008552

7566127

2 (bases 1 to 5989)

Chen, J.D. and Evans, R.M.

Direct Submission

Submitted (27-SEP-1995) J. Don Chen, Gene Expression Lab, The Salk Institute, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

location/Qualifiers

1..5989

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_line="HeLa"

1..5989

/gene="SMRT"

496..4983

/gene="SMRT"

/note="transcriptional co-repressor"

/codon_start=1

/product="silencing mediator of retinoid and thyroid hormone action"

/protein_id="AAC50236.1"

/db_xref="GI:1045655"

/translation="MEAWDAHPDKEAFABAQKLPQDPPCWTSLGLPPPPPREVIKAS PHAPDGFASYAPPGHPLPLGLHPTARVLPVLPPTISNPPPLISSAKHPSVLEIQA ISQMGVQLHVYSEHAKAPVGMGLPMDPKLAPFSGVKQLSPRGQAGPYE SLGVPYAOEASVLRTGALSGVPGGSI TKGI PSTVPSDSALT YRGSITHTGTADLYK GPTRIIIGEDSPRLDRGSDLPKHVIYVEGKGHVLSEYSGMSVTVQCSKEDGRSSS GPHETAPKRTYDWMGRVGRALSSAS IEGLMGRAIPPERHSPHLKQKHIRGSIT QIPRVSVEAOYDLRBAKLLKEGTGTPPPPSDLTEAYKTOALGPKLKEAHEGL VATVKGASISHEIPRELRHTPELPLAPLAKESITCGTPKIDTGASTTGSKGRHV VRLSGIPKTPPPVHPDWDARALERACEESLKSPGTSASSGSGSAGKAPVIV PELKSPQSLTYEDHGFAGHLPGRSPVPMKEPRLQEGSSUSSKASQQRKTUT PREAKSPSTVEHPHPHIFSYELLHGVSGVDLYRSHIPLADPTSPRGPIDAA AAYLPRHLAPNPTPHLYPPYLRGYPDPALENQRTIINDYITSQKHNTATAMA QYADMLGSPRESSLALAGPRGIDLSQVHPLVLPVPTTEGTATAMDRLAYLP TAPQPTSSRHSPLSPGSPHPLTKPTTSSSERDRDRDRDREREKSLITSTTT VEHAPITWRPQTQSGSGSGSGSGSSSRPASHAHQHSPIPSPTQDALQORPSVL

exon

ORIGIN

Alignment Scores:

Pred. No.: 978e-152 Length: 5989
Score: 7885.00 Matches: 1510
Percent Similarity: 91.34% Conservative: 7
Best Local Similarity: 91.91% Mismatches: 14
Query Match: 51.67% Indels: 77
DB: 9 Gaps: 7

US-09-522-753-5 (1-2517) x HSU37146 (1-5989)

Qy 984 ProProAspGluAspAlaAlaProThrLysProAlaProProAla 998
Db 168 CCACCAAGA-----TCACCATATGAGCCACGACCCCTGCATAGGAAATATCATAT 218
Qy 999 ProProProPro----- 1002
Db 219 CCGGAGCCACCTGTGTTTCATTAAGCCATCTGCCATTCGTATGAGGCTCCATGTATCTATT 278
Qy 1003 -----GlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGly 1016
Db 279 GCAGAGTCCCAATTCATCAGCACTGAAACCAACCATCGCTCGTAACGGTCTCCAGT 338
Qy 1017 -----SerSerPro-----Arg 1020
Db 339 CTCCCTCTCTCTCATAGGCCATGAGGCTCTCCCTCTAGGTGGTGGTGGAGGAGA 398
Qy 1021 GlyLysSerArgSer-----ProAlaProPro-Ala----- 1030
Db 399 GGAAGATTCCGAGCTCTGTACCAACCCCGCGCGCTCGTCCGGAGAGGAGGAGTGGT 458
Qy 1030 ----- 1030
Db 459 CCTCGACGAGGCTCATATCATATAATCTCTTAGATGAGGATGGAGCCACCC 518
Qy 1031 -AspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProCysTrpTh 1050
Db 519 CGACAGAGAGGCTTCACGACCGAGCCGAGCTGCTGGGAGCCCCCTTGTCGAC 578
Qy 1050 rSerGlyLeuPheProValProProArgGluValLysAlaSerProHisAlaPr 1070
Db 579 TTCGGGCTGCTCTTCCCGCTGCCCGCGCTGAGGTGATCAAGGCTCCCGCATGCC 638
Qy 1070 oAspProSerAlaPheSerTyraProProGlyHisProLeuProLeuGlyLeuHis 1090
Db 639 GGACCCCTCAGCTTCTCTTCCCTAGCTCCACCTGGTCCACCATCTGCCCTCCATGA 698
Qy 1090 pThrAlaArgProValLeuProArgProProThrLysAsnProProProLeuLeSe 1110
Db 699 CACTCCCGCGGCTCTCTCCCGCGCCACCCACCATCTCCACCCCTCCCTCATCTC 758
Qy 1110 rSerAlaLysHisProSerValLeuGluArgGlnLysAlaLysSerGlnLysMetSe 1130
Db 759 CTCTGCCAAGCTCCCAAGCTCTCCGAGGAGCAATAGTGCCATCTCCCAAGGAATGTC 818
Qy 1130 rValGlnLysHisValProTySerGluHisAlaLysAlaProValGlyProValThrMe 1150

Db 819 GGTCCAGCTCCAGCTCCCTGCTCAGAGCATGCAAGGCCCGGTGGGCTGCTGACCAT 878
Qy 1150 tGlyLeuProLeuProMetAspProLySylsLeuAlaProPheSerGlyVallyslngl 1170
Db 879 GGGGCTGCTCCCTCCATGACCCCAAAAAGCTGGCACCTTCAGCGGAGTGAAGCAGA 938
Qy 1170 uGlnLeuSerProArgGlyGlnAlaGlyProProGlnSerLeuGlyValProThrAlaG 1190
Db 939 GCAGCTGTCCCAAGGGCCAGGCTGGGCCACCGAGAGCCCTGGGGGTGGCCACACCC 998
Qy 1190 nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThr 1210
Db 999 GAGAGCGCTCGCTGAGAGGACAGCTCTGGGCTCAGTTCGGGCGGAAAGCATCAAC 1058
Qy 1210 sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyArgGlySerIleTh 1230
Db 1059 AGGCATTCCTCCAGCACACGGGTGCTCGGACAGCGCATCATACACCGCGCTCC 1118
Qy 1230 rHisGlyThrProAlaAspValLeuTyLySylsGlyThrIleThrArgIleLeGlyGlu 1250
Db 1119 CCAGCGCACCCAGCTGACGCTCTGTACAGGGGACATCACAGAGATCATCGCGAGA 1178
Qy 1250 pSerProSerArgLeuAspArgGlyA:ggluAspSerLeuProLySylsValIleTh 1270
Db 1179 CAGCCCGAGTGGCTTGGACCGCGCGGAGGACAGCTCTGCCCAAGGCGCCACGTCAT 1238
Qy 1270 rGluGlyLySylsGlyHisValLeuSerTyArgGlyGlyMetSerValThrGlnCys 1290
Db 1239 CGAAGGCAAGAGGGCCAGCTCTGTCTATGAGGTGGCATGTCTGTGACCCAGTGCTC 1298
Qy 1290 rLySylsGlyArgSerSerGlyProProHisGluThrAlaAlaProLySylsArgTh 1310
Db 1299 CAAGGAGGAGCGGAGAGCAGCTCAGGACCCCGCCCATGAGACGGCGCGCCCAAGGCG 1358
Qy 1310 rTyArgMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGlyLe 1330
Db 1359 CTATGACATGATGAGGGCGCGTGGGAGAGGACATCTCTCAGCCAGCATCGAAGTCT 1418
Qy 1330 uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLySylsGluGln 1350
Db 1419 CATGGCGGTGCCATCCCGCGAGGACACAGCCCGCCCATCAGCTCAAGAGGACGAC 1478
Qy 1350 sIleArgGlySerIleThrGlnGlyIleProArgSerTyValGluAlaGlnGluApy 1370
Db 1479 CATCCGGGGTCCATCACAAAGGATCCCTCGCTCTACGTGGAGGACAGGAGACTA 1538
Qy 1370 rLeuArgArgGluAlaLySylsLeuLySylsArgGlyThrProProProProProSe 1390
Db 1539 CCTGCTCGGGAGGCCAAGCTCTCAAGGGGAGGACAGCGCTCGCGCCCGCCACCGCTC 1598
Qy 1390 rArgAspLeuThrGluAlaTyLySylsThrGlnAlaLeuGlyProLeuLySylsProAl 1410
Db 1599 ACGGAGCTCAGCGAGGCTTACAGAGCGAGGCGCTGGGCGCCCTGAGAGTGAAGCCGC 1658
Qy 1410 aHisGluGlyLeuValAlaThrValLySylsGluAlaGlyArgSerIleHisGluIlePro 1430
Db 1659 CCATGAGGGCTGCTGGCCACCGGTGAAGGAGGGCGGCGCTCCATCCATGAGATCCCG 1718
Qy 1430 gGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLySylsGly 1450
Db 1719 CGAGAGCTGGGACACAGCCCGAGCTGCGCTGGCGCCCGCGCGCTCAAGGAGGCTC 1778
Qy 1450 rIleThrGlnGlyThrProLeuLySylsThrArgThrGlyAlaSerThrThrGlySer 1470
Db 1779 CATCAGCGAGGACCGCGCTCAAGTACGACACCGGGGCTCCACCACTGGCTCCAAAA 1838
Qy 1470 sHisAspValArgSerLeuIleGlySerProGlyArgThrPheProValHisProLe 1490
Db 1839 GCAGAGCTACGCTCCCTCATCGGACAGCCCGCGCGGAGGCTTCCACCGCTGACCGCT 1898
Qy 1490 uAspValMetAlaAspAlaLeuGluArgAlaCysTyArgGluGlnSerLeuLySyl 1510
Db 1899 GATGTGATGGCGAGCGCCCGGCACTGGAACTGCTGCTACGAGGAGAGGAGCTTGAGAG 1958

1510 rArgProGlyThrAlaSerSerGlyGlySerIleAlaArgGlyValaProValIleVa 1530
1599 CCGGCCAGGAGAGCCAGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGCGCCCGCTCATGT 2018
1530 lProGluLeuGlyArgProArgGlnSerProLeuThrTyArgAspHisGlyAlaProPh 1550
2019 GCTGAGCTGGTAAGCCCGGAGAGCCCTGACCTATGAGGACCAACCGGACCCCTT 2078
1550 eAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuG 1570
2079 TCCCGGCGCACCTCCACAGAGTTCCCGGTGACCATGCGGAGCCAGCGCGCGCTGCA 2138
1570 nGluGlySerLeuSerSerLySylsAlaSerGlnAspArgLySylsLeuThrSerThrPro 1590
2139 GAGGGCAGCGCTTCGTCCAGCAAGGCATCCAGGACCGAAAGCTGAGTGCAGCGCTCG 2198
1590 gGluIleAlaLeuSerProHisSerThrValProGluHisHisProHisProIleSerPr 1610
2199 TGGATCGCCAGCTCCCGGACAGCACCGTCCCGGAGCACCACCCACCCCATCTGCC 2258
1610 oTyroGluHisLeuLeuArgGlyValSerGlyValAspLeuTyArgSerHisIleProLe 1630
2259 CTATGAGCACCTTCCTCGGGGCGTGAAGTGGCGTGACCTGTATCGCGAGCCACATCCCGCT 2318
1630 uAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaTy 1650
2319 GGCTTCGACCTACCTCCATACCCCGCGCATCCCTCTGGACGCGAGCGCTGCTACTA 2378
1650 rLeuProArgHisLeuAlaProAspProThrTyProHisLeuTyProProTyLeuI 1670
2379 CTGCGCGGAGAGCTGGCGCCCAACCCACCTACCCGACCTGACCTACCCACCTACCTCAT 2438
1670 eArgGlyTyProAspThrAlaAlaLeuGluAspArgGlnThrIleLeuAspTyI 1690
2439 CCGCGGTACCTGACACGCGCGCGCTGGAGAACCGGACGAGCATCATCAATGACTACAT 2498
1690 eThrSerGlnGlyMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLe 1710
2499 CACTCGCAGCGATGACACCAACAACAGCCCGCCATGGCCGAGCGAGCTGATGCT 2558
1710 uArgGlyLeuSerProArgGlnSerSerIleAlaLeuAsnTyArgAlaGlyProArgG 1730
2559 GAGGGGCTCTGCGCGCGGAGTCTCTCGTGGCACTCACTACGCTGGGGTCCCGGAGG 2618
1730 yIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyTh 1750
2619 CATCATCGACCTTCGCAAGTGCCACACCTGCTGTGCTCGTCCCGCCGACACCGAGGAC 2678
1750 rProAlaThrAlaMetAspArgLeuAlaTyLeuProThrAlaProGlnProPheSerSe 1770
2679 CCGACGACCGATGAGACCGCTTGCCTACTCCACCGCGCGCCCGCCCTTCAGCAG 2738
1770 rArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLySylsProThr 1790
2739 CCGCCACAGCAGCTCCCTCCCTCTCCAGGAGTCCCAACACACTTGACAAACCAACAC 2798
1790 rThrSerSerSerGluArgGluArgAspArgGluArgAspArgAspArgGluAr 1810
2799 CAGCTCTCTGTGAGCGGAGGACAGACCGGGATCGAGAGCGGAGCCGCGATCGGAGCG 2858
1810 gGluLySerIleLeuThrSerThrThrValGluHisAlaProIleTTPArgProG 1830
2859 GGAAGTCCATCTCATGTCACACCGAGCGGTGGAGCAGCAGCCCATCTCGAGACCTGG 2918
1830 yThrGluGlnSerSerGlySerSerGlySerGlyGlyGlyGlyGlyGlySerSerAr 1850
2919 TACAGAGCAGCAGGAGCGGAGCGGAGCGGCGGGGTGGGGGCGAGCAGCAGCCG 2978
1850 gProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAl 1870
2979 CCGCGCTCCCTTCCTCCATCCAGCCACGACGACTGCGCCCATCTCCCTCGGACCCAGGATGC 3038

QY 1870 aLeuGlnGlnArgProSerValLeuHisLeuThrGlyMetLeuGlyLeuThrAlaVal 1890
DB 3039 CTTCCAGCAGACCCAGTGTCTTCAACACAGGATGAAGGGTATCATCACCGCTGT 3098
QY 1890 lGluProSerLysProThrValLeuArgSerThrSerSerProValArgProAl 1910
DB 3099 GGAGCCAGCAGACCCAGTGTCTTCAACACAGGATGAAGGGTATCATCACCGCTGT 3158
QY 1910 aAlaThrPheProAlaThrHisCysProLeuGlnGlyLeuThrLeuAspGlyValTyPr 1930
DB 3159 TGCCATTTCCCACTGCCACCCACTGCCACTGGGCGGCAACCTCGATGGGCTTACCC 3218
QY 1930 oThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluAr 1950
DB 3219 TACCTCTCATGGAGCCCGTCTTGTGCTCCAGAGAGGCCCGCCCGGCTGCCCGCCAGAGCG 3278
QY 1950 gProArgAlaAspThrGlyHisAlaPheLeuAlaLysProAlaArgSerGlyLeuG1 1970
DB 3279 GCCCGGAGCAGACCCGCTTCTTCTGCCCAAGCCCGCCAGCCGCTGCCGGCTGGA 3338
QY 1970 uProAlaSerSerProSerLysGlySerGluProArgProLeuValProProValSerG1 1990
DB 3339 GCCCGCTCTCTCCCGCAGAGGGCTCGGAGCCCGCCCTTGTGCTTCTCTCTCTCTG 3398
QY 1990 yHisAlaThrLeuAlaArgThrProAlaLysAsnLeuAlaProHisAlaSerProAs 2010
DB 3399 CCACGCCACCATCGCCGCCACCCCTCGCAAGAACCTCGCACCTCACCAAGCCCGCGGA 3458
QY 2010 pProProAlaProProAlaSerAlaSerAspProHisArgGluLeuThrGlnSerLysPr 2030
DB 3459 CCGCGCGCGCCACTGCTGCTCGGCTCGGACCGCCGCGGGAAGACTCAAGTAAACC 3518
QY 2030 oPheSerLeuGlnGluLeuArgSerLeuGlyTyHisGlySerSerTySerPr 2050
DB 3519 CTTTTCATPCAGAGACTGGAACTCCGTTCTCTGGGTTTACCGGCGAGCAGCTACAGCCC 3578
QY 2050 oGluGlyValGluProValSerProValSerProSerLeuThrHisAspLysGlyLe 2070
DB 3579 CGAAGGGTGGAGCCGCTGAGCCCTGTGAGCTCACCCAGTCTGACCCACCAAGGGGCT 3638
QY 2070 uProLysHisLeuGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysG1 2090
DB 3639 CCCCAGCACCTCGAAGAGCTCGACAGAGCCACTCGAGGGGGAGCTGCGGCCAGCA 3698
QY 2090 nProGlyProValLysLeuGlyGlyAlaAlaHisLeuProHisLeuArgProLeuPr 2110
DB 3699 GCCAGGCCCGTGAAGCTTGGCGGGAGCGCCGACCTCCACACTGCGCGCGCTGCC 3758
QY 2110 oGluSerGlnProSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisG1 2130
DB 3759 TGAGAGCCAGCCCTGCTGTCAGCCGCTGCTCCAGACCGCCCGCCAGGGGTCAAAGTCA 3818
QY 2130 nArgValValThrLeuAlaGlnHisLeuSerGluValLeuThrGlnAspTyThrArgH1 2150
DB 3819 GGGGTGTCTCACTTGGCCAGCAGCATAGTGGAGTCTATCACAGAGACTACCCGCGCA 3878
QY 2150 shiisProGlnGlnLeuSerAlaProLeuProAlaProLeuTySerPheProGlyAlase 2170
DB 3879 CCACCCAGCAGCAGCTCAGCGCACCCCTGCGCGCCCTCTCTCTCTCTCTCTCTCTCT 3938
QY 2170 rCysProValLeuAspLeuArgProProSerAspLeuTyLeuProProProAspH1 2190
DB 3939 CTGCCCCGTCTGAGACTTCCGCGCCGCCACCCAGTGACCTCTACCTCCCGCCCGCGACCA 3998
QY 2190 sGlyAlaProAlaArgGlySerProHisSerGluGlyLysArgSerProGluProAs 2210
DB 3999 TGTGTCCCCCGCTGTGTCTCCCGCAGGAGGGGCGAGAGGTCTCCAGAGCCAAA 4058
QY 2210 nLysThrSerValLeuGlyGlyGlyGluAspGlyLeuGluProValSerProProGluG1 2230
DB 4059 CAAAGACGTGCGTCTTGGGTGGTGGAGACGGTATTGAACCTGTGTCCCGCCAGGAGG 4118
QY 2230 yMetThrGluProGlyHisSerArgSerAlaValTyProLeuLeuTyArgAspGlyG1 2250

DB 4119 CATGACGGAGCTAGGGCACTCCCGGAGTGTGTGTACCGGCTGTGTACCGGATGGGA 4178
QY 2250 uGlnThrGluProSerArgMetGlySerLysSerProGlyLeuThrSerGlnProProAl 2270
DB 4179 ACAGACGGAGCTCAGCAGGATGGCTCCAACTCTCCAGGCAACACCGCCAGCGCGCAG 4238
QY 2270 aPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGlu1 2290
DB 4239 CTTCTTCAGCACTGTGACCGAGAGCACTCCCGCATGGTCAAGTCCAGAGCAAGAGAT 4298
QY 2290 eAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyArgLysLeuSerGlnProG1 2310
DB 4299 CAACAGAAAGCTCAACACCCAGATGAGCTGAATACATATATCAGCAGCGCTGG 4358
QY 2310 YThrGluLeuPheAsnMetProAlaLeuThrGlyLeuMetThrTyArgSerG1 2330
DB 4359 GACGGAGATCTTCAATATGCCCGCATCACCGGACAGGCTTTATGACCTTATAGAGCCA 4418
QY 2330 nAlaValGlnGlnHisAlaSerThrAsnMetGlyLeuGluAlaLeuLeuArgLysAlaLe 2350
DB 4419 GGCGGTGCGAGGACATGCCAGCACCAATGGGGCTGGAGGCCATAATTAGAAGCACT 4478
QY 2350 uMetGlyLysTyAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnPr 2370
DB 4479 CATGGGTAAATATGACCAAGTGGGAAGAGTCCCGCGCTCAGCGCCAACTGCTTTAACCC 4538
QY 2370 oLeuAsnAlaSerAlaSerLeuProAlaAlaMetProLeuThrAlaAlaAspGlyArgSe 2390
DB 4539 TCTGATATGCATGTCAGCTGCGCTGCTATGCTTATGCTTATGCTTATGCTTATGCTT 4598
QY 2390 rAspHisThrLeuThrSerProGlyGlyGlyLysAlaLysValSerGlyArgProSe 2410
DB 4599 TGACCAACACACTCACCTCGCAGGTGGCGCGGAGGCCCAAGGTCTCTGGCAGACCCAG 4658
QY 2410 rSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgProProSerVa 2430
DB 4659 CAGCCGAAAGAGTCAAGTCCCGCGCTGGCACTCTGGGACCGCGCCACCTCTGT 4718
QY 2430 lSerSerValHisSerGluGlyAspCysAsnArgThrProLeuThrAsnArgValTr 2450
DB 4719 CTCTCATGTGCTCTCGAGGGAGACTGCAACCGCGGAGCGCTTACCAACCGCGGTG 4778
QY 2450 pGluAspArgProSerSerAlaGlySerThrProPheProTyArgProLeuLeuMetAr 2470
DB 4779 GSAGGACAGGCTTCTGTCGCGAGGTTCCACGCACTTCCCTTACAAACCCCTGATATGCG 4838
QY 2470 gLeuGlnAlaGlyValMetAlaSerProProProGlyLeuProAlaGlySerGlyPr 2490
DB 4839 GCTGACGCGGCTGTGATGCTTCCCGCCCGCGGCTTCCCGCGGCGCGCGGCGCC 4898
QY 2490 oLeuAlaGlyPheHisAlaTrpAspGluProLysProLeuLeuCysSerGlyTy 2510
DB 4899 CTTGCTGCGCTTACACACGCTTGGGAGGAGGCCCAAGCCACTGCTTCTGTCTGCGAGTA 4958
QY 2510 rGluThrLeuSerAspSerGlu 2517
DB 4959 CGAGACACTCTCGACAGCGAG 4980

RESULT 7

S83390

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

2930 bp mRNA linear PRI 28-MAR-1997
T3 receptor-associating cofactor-1 [human, fetal liver, mRNA, 2930
nt]

S83390

S83390.1 GI:1911769

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa;

Mammalia; Eutheria;

Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 2930)

Sande, S. and Privalsky, M.L.

